PRENATAL TESTOSTERONE EXPOSURE AND NUMERICAL COMPETENCE IN CHILDREN AND ADULTS

HELEN BROOKES

PhD

2010

Prenatal testosterone exposure and numerical competence in children and adults

HELEN BROOKES

A thesis submitted in partial fulfilment of the requirement
of Northumbria University for the degree of
Doctor of Philosophy

Research undertaken in the School of Psychology &
Sport Sciences

April 2010

Abstract

The present thesis sought to investigate the potential relationship between the second to fourth finger ratio (2D:4D), as a somatic marker of prenatal testosterone exposure, and basic numerical skills in children and adults. Chapter 1 presents a basic overview of the nature and effects of sex steroids followed by a more comprehensive consideration of literature regarding the reported effects of prenatal testosterone (PT) on the brain and cognition. The chapter then more specifically considers the possible influence of PT on numerical and mathematical competencies.

Experiment 1 attempted to replicate evidence for a relationship between 2D:4D and basic numerical skills in children. The results revealed only one significant correlation, namely a significant positive correlation between right hand 2D:4D and number comparison scores in females.

Chapter 3 discussed research regarding the nature and characteristics of so called 'core' numerical competencies. Experiments 2-4 then attempted to explore any relationship between 2D:4D and performance on tasks designed to assess such skills in adults. The results of all three studies revealed an association between 2D:4D and lateralization for the process of subitizing relative to a comparable control task. The nature of this observed effect however varied across the three experiments. Experiment 4 also identified significant positive correlations between left hand 2D:4D and counting reaction times in females and a series of two way interaction effects between 2D:4D and task (numerical vs. control) for subitizing, counting and number comparison performance. The revealed interactions predominantly suggested faster task reaction times/higher accuracy in high 2D:4D (low PT) participants as compared to low 2D:4D (high PT) participants on the numerical tasks and the opposite pattern of results (i.e. high 2D:4D associated with poorer performance) on the control tasks.

Experiment 5 investigated the association between 2D:4D and core numerical skills in children. Significant correlations were observed between; left hand 2D:4D and subitizing reaction times to the left visual field in males (negative direction), right hand 2D:4D and subitizing reaction times the right visual field in females (positive direction) and left hand 2D:4D and subitizing percentage error scores to the right visual field in females (negative direction). A possible relationship between 2D:4D and lateralization for both subitizing and number comparison relative to control was also found. For both numerical tasks low 2D:4D participants showed a right visual field advantage and high

2D:4D participants showed a left visual field advantage while different patterns of results were shown on the control task.

Experiment 6 re-considered the relationship between 2D:4D and basic and core numerical skills in children using a standardised assessment of numerical competencies (the Dyscalculia Screener). No significant correlations however between 2D:4D and performance were identified.

Finally, experiment 7 re-examined evidence for a link between 2D:4D and Key Stage 1 Standardised Assessment Test (SAT) scores. While the findings did not replicate evidence for a direct link between 2D:4D and SAT numeracy scores the results did demonstrate a significant negative relationship between right hand 2D:4D minus left hand 2D:4D (Dr-1; higher scores thought to indicate lower exposure to PT) and SAT numeracy scores in females. Such findings may potentially suggest a facilitative influence of PT on numeracy in women.

Overall, while a number of interesting findings were revealed, limited consistency was identified across the results of the experiments conducted in the present thesis. The findings therefore offer no concrete support for a possible association between 2D:4D and basic numerical skills in either children or adults. The final chapter summarises the findings of each experiment and considers the results in the context of previous literature. General limitations of the research and suggestions for future research are discussed.

Contents

Acknowledgements

A huge thank you to my principal supervisor Nick Neave and my second supervisor Colin Hamilton for their ongoing support and encouragement throughout this research programme. To my parents and my fiancé Neil, your ongoing love, support, encouragement and guidance has kept me going, thank you for everything.

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the School Ethics Committee.

Name: Helen Brookes

Signature:

Date:

Chapter 1

Introduction

1.1 The endocrine system

The endocrine system is one of the major modes of intercellular and inter-organ communication within the body. Hormones are produced and released by specialized secretary cells in the endocrine glands and carried via the bloodstream, to target cells, where they bind to specific receptors to exert their effects (Hardie, 1991). As hormones can travel through the blood to virtually every cell in the body, they can potentially impact upon any organ or cell containing appropriate receptors (Nelson, 2000). The effects of hormones are broad, and extend across the life span, beginning early in the fetal period (Kawata, 1995; Keenan, & Soleymani, 2001). In total, the body creates more than 50 hormones performing a range of functions including; maintaining metabolism, mediating the use and storage of energy, regulating the rates of chemical reactions in cells, and promoting growth (Erlanger, Kutner, & Jacobs, 1999).

The brain plays a fundamental role in the mediation and initiation of hormone production and release. Chemical communication between the hypothalamus and pituitary gland directly regulates, via positive or negative feedback mechanisms, the majority of endocrine function and levels of hormone concentration in the body. Crucially however, the relationship between the brain and the endocrine system is now seen to extend far beyond the regulation of somatic hormone production by the hypothalamus and pituitary. The brain itself can be considered both as an endocrine organ, producing hormones that act both within and outside the central nervous system, and, as an important target hormone for endocrine effects (Gooren, 2007). Burgeoning knowledge regarding the pervasive impact of hormones has led to a concomitant accumulating interest in the endocrine system as one source of individual differences in physiological, behavioural, and cognitive development. One of the most widely studied classes of hormones with regard to their potential effects on neural anatomy and subsequent behavioural and cognitive outcomes are the sex steroid hormones.

1.2 Sex Steroid Hormones

The sex (or gonadal) steroids include the androgens, estrogens, progestins and their metabolites. The main androgen hormones are dehydroepiandrosterone (DHEA), androstenedione, androstenediol, androsterone and dihydrotestosterone (DHT), with the primary and most widely-known androgen being testosterone (T). Androgens perform a number of functions in the human body including: the differentiation and maturation of the male reproductive organs, the initiation and maintenance of spermatogenesis, the promotion of sexual maturity at puberty, the development of male secondary sexual characteristics (e.g. beard growth), and the control of sexually dimorphic reproductive behaviour patterns (Baron-Cohen, Luchmaya, & Knickmeyer, 2004; Norman & Litwack, 1997). There are three major naturally occurring estrogens, namely, estradiol, estrone, and estriol. All three are involved in the maintenance of pregnancy and the development of female secondary sexual characteristics. Estrogens also play an important role in the prevention of bone mineral loss and the distribution of fat on certain body regions (Baron-Cohen et al., 2004). Progestins are typically considered to be antiandrogens, but can demonstrate some androgenic effects (Clark, Schrader, & O'Malley, 1985; Hendricks, 1992). The principle progestin is progesterone which plays a vital role in the preparation of the uterus for implantation of the ovum, pregnancy maintenance, and preparation of the mammary glands for lactation. Although estrogens and progestins are usually considered to be 'female' hormones, and androgens 'male' hormones, the gonadal hormones are not sex specific, but are secreted by both males and females, although in hugely different amounts. Thus the sexes differ with regard to both the quantity of each gonadal hormone present and the number of receptors for them (Baron-Cohen et al., 2004).

Sex steroid hormones vary widely during key growth phases across the life span: in the fetal and perinatal environments, during adolescence, and as a natural consequence of aging (Erlanger et al., 1999). Males experience a surge in testosterone concentration at approximately weeks 8-24 of gestation and again in months 1-5 of postnatal life (Cohen-Bendahan, van de Beek, & Berenbaum, 2005; Smail, Reyes, Winter, & Faiman, 1981). Females experience a postnatal surge of estrogen, with high levels of exposure maintained throughout the first year of life, peaking around months 3-4 (Bidlingmaier, Strom, Dorr, Eisenmenger, & Knorr, 1987). Following this postnatal surge, sex steroid levels in both males and female throughout childhood remain relatively low (Chada et al., 2003; Ducharme, Forest, De Peretti, Sempe, & Bertrand, 1976; Erlanger et al., 1999). With the onset of puberty, males and females experience a secondary postnatal surge in sex hormone concentrations. Elevated levels of sex steroid exposure during this period are vital for the development of secondary sexual characteristics and the overt signs of reproductive maturation such as breast development. Following the onset of puberty, serum levels of testosterone are reported to show a marked sexual dimorphism with higher concentrations identified in males (approx. 300-1000 ng/dl.) as compared to females (approx. 30-70ng/dl) (Erlanger et al., 1999). Serum levels of estrogen in females vary widely according to the menstrual cycle (Erlanger et al., 1990; Norman & Litwack, 1987). Levels of estradiol, for example, are reported by Erlanger et al. (1990) to range from approximately 20-100 pg/ml in the early follicular phase to 100-350 pg/ml in the preovulatory and luteal phases. Levels of serum estradiol in adult males are reported to range from 20-50 pg/ml (Erlanger et al., 1999).

1.3 The effects of sex steroids

The effects of sex steroids have traditionally been classified into two distinct categories, i.e. organisational and activational. Organisational effects refer to those that are permanent and result in changes in the way the brain is organised, i.e. its structural characteristics. The time frame for organisational influences is usually hypothesised to occur early in development, coinciding with the male fetal testosterone peak during weeks 8-24 (Cohen-Bendahan et al., 2005; Collaer & Hines, 1995). Early postnatal surges in gonadal hormones may also constitute another sensitive period during which hormones may exert an organising impact. The precise physiological and neural effects of elevated sex steroid exposure in humans during the early postnatal period however remain poorly understood. Notably, while it has been generally assumed that activational effects occur during adulthood, whereas organisational effects occur during early development (Buchanan, Eccles, & Becker, 1992), an expanding body of literature now suggests that adolescence may also constitute a period of development during which sex hormones can exert an organisational effect on the nervous system (Romeo, 2003; 2005; Romeo, Richardson & Sisk, 2002; Sisk, Schulz & Zehr, 2003; Sisk & Zehr, 2005).

Activational effects refer to those that occur later in development as a result of circulating hormone levels, are superimposed on early organisational effects, and result in transient, time-limited, functional changes in neural circuitry (Knickmeyer & Baron-Cohen, 2006; Walker, Sabuwalla & Huot, 2004). Research suggests that activational effects are often essential in order to allow the tissue or organ in question to perform its function. For example, evidence in animals shows that while the tissues of the genetic male are organised pre-natally for adult reproductive behaviour, such behaviour may not be displayed in the absence of appropriate adult sex hormone exposure (Knickmeyer & Baron-Cohen, 2006; Phoenix, Goy, Gerall, & Young, 1959).

1.4 Sex steroids and cognition

According to Arnold (1996) virtually every neural sexual difference studied in animals is the result of documented sexually dimorphic sex steroid secretion (see discussion below). As it is assumed that the mechanism of hormone action in the animal and human brain are broadly very similar, it is hypothesised that gonadal hormones may also account for one mechanism of cognitive sexual differentiation in humans.

While no sex differences are reported with regard to overall global intelligence, evidence does suggest that differences between males and female exist in patterns of cognitive performance. There is evidence that men outperform women in certain tasks assessing spatial orientation and visualization, mechanical knowledge and mathematical reasoning, whereas women outperform men in certain tasks assessing verbal fluency, verbal learning, verbal memory, emotional perception, fine motor skill, and perceptual speed (Halpern, 2000; Hamilton, 2008; Kimura, 1999; Maccoby & Jacklin, 1974). As all human societies attribute major significance to the status of being male or female, sex differences have been the subject of both general public concern and scientific interest since the origin of cognitive research. The possibility that some variation in cognitive performance may be related to inter-individual differences in chronic levels of gonadal hormones has made the neural and cognitive impact of androgens and estrogens a popular focus for scientific investigation. As foetal androgens, in particular testosterone, play a major role in prenatal stages of sexual differentiation, the organising effects of testosterone have, in particular, been the subject of considerable research attention. This thesis focuses upon the potential prenatal organising effects of testosterone upon certain aspects of cognition.

1.5 Testosterone

Like all steroid hormones the precursor to testosterone is cholesterol. Testosterone production is initiated when cholesterol is converted to pregnenolone. There are two possible metabolic routes that then lead from pregnenolone to testosterone, namely the Δ^5- (testosterone from pregnenolone via 17α-hydroxypregnenolone, dehydroepiandrosterone and androstenedione) pathway and the Δ^4- (testosterone from pregnenolone via progesterone, 17α-hydroxyprogesterone and androstenedione) pathway (Ohno, Nakajima & Nakajin, 2005). Testosterone itself can be converted to estrogens via a process named aromatization. Testosterone is furthermore a precursor to the potent non-aromatizable androgen 5α-dyhydrotestosterone (DHT). In males, testosterone is primarily produced in the Leydig cells of the testes with small amounts also derived from the adrenal glands. In females it is estimated that both the theca interna cells of the ovary and the adrenal cortex each contribute to approximately 50% of the plasma levels of testosterone (Palacios, 2007).

1.5.1 Circulating testosterone

There is now a large body of evidence demonstrating certain activational effects of testosterone (i.e. phasic influences resulting from fluctuations in circulating hormone levels) on some behavioural and cognitive measures. The precise nature of this potential relationship however remains controversial. A review of current research on the relationship between circulating testosterone levels as determined via blood or saliva immunoassay for example reveals evidence for, *1)* a curvilinear relationship (whereby medium concentrations of androgens may facilitate spatial performance whilst higher and lower levels of testosterone may inhibit performance) between circulating testosterone and spatial performance (Gouchie & Kimura, 1991; Moffat & Hampson, 1996; Neave, Menaged, & Weightman, 1999; Shute, Pellegrino, Hubert, & Reynolds, 1983), *2)* a positive linear relationship between circulating testosterone and spatial performance (Christiansen, 1993; Christiansen & Knussmann, 1987; Gordon & Lee, 1986), and *3)* no relationship between the two factors (McKeever, Rich, Deyo, & Conner, 1987; McKeever & Deyo, 1990). Evidence regarding a potential relationship between circulating testosterone level (measured via blood or saliva) and verbal cognition is equally mixed. While evidence from Christiansen and Knussmann (1987) and Christiansen (1993) reports a significant negative correlation between circulating testosterone and verbal abilities research by Gouchie and Kimura (1991), Moffat and Hamson (1996), and Neave et al. (1999) found no significant relationship between circulating testosterone and verbal task performance. Beyond spatial and verbal skills, Gouchie and Kimura (1991) also report a significant curvilinear relationship between levels of circulating testosterone and measures of mathematical ability. There is also some evidence for a positive relationship between measures of salivary testosterone and simple reaction time (Müller, 1994).

As the current project focuses on the organising effects of testosterone an in depth review of the empirical evidence regarding the association between circulating testosterone and cognitive performance is beyond the scope of this thesis. It is vital to recognise however that research regarding the influence of testosterone is not limited to its potential organising effects.

1.5.1 Prenatal testosterone

Acting prenatally, testosterone plays an integral role in the development and masculinisation of the male reproductive organs, thus while sex is determined genetically, sexual differentiation is thought to be largely hormonal. The human embryo is initially bi-potential such that it possesses bipotential gonads that resemble neither testes nor ovaries. Prior to sexual differentiation both XX (genetically female) and XY (genetically male) individuals have two sets of ducts connecting the indifferent gonad to the exterior: the Müllerian (female) ducts and the Wolffian (male) ducts. Whether the fetus will develop testis or ovaries is determined by the presence of a Y-linked gene referred to as the sex-determining region of Y chromosome (SRY). The SRY causes the development of testes. In the absence of this gene the gonads will develop as ovaries. The development of the testes results in the concomitant release of testosterone and a second hormone called Müllerian Inhibiting Hormone (MIH). MIH induces the regression of the Müllerian ducts in turn preventing the development of female reproductive organs while testosterone induces the differentiation of the Wolffian ducts and the transforms the genitals into male organs. As females have no Y chromosome, the fetal ovaries do not produce MIH during development creating a permissive environment for differentiation of the Müllerian ducts. The female reproductive system therefore develops by default. The absence of testosterone or functioning receptors leads to the passive regression of the Wolfian duct (male reproductive) system. Testosterone therefore is generally seen as the critical factor in the sexual differentiation of the fetus via masculinisation. Feminization is typically viewed as a passive process that occurs in the absence of high levels of androgens (although a number of authors have highlighted the importance of ovarian hormones for complete feminization, e.g. Dimond, Dowling, & Johnson, 1981; Dohler et al., 1984; Fitch, Cowell, Schrott, & Denenberg, 1991; Leret, Monlina-Holgado, & Gonzales, 1994; Stewart & Cygan, 1980).

As well as playing a key role in the differentiation of the gonads and reproductive system testosterone has an indisputable effect on the brain and behaviour. In a seminal paper, Phoenix et al. (1959) reported that female guinea pigs exposed to testosterone during gestation showed more masculinised and less femininised copulatory behaviours if again provided with testosterone as adults. Female fetuses exposed to androgens as adults but not during development, showed no such effects. Since this pioneering study a huge amount of research utilising the experimental manipulation of testosterone in animals has confirmed that the effects of testosterone during pre- and peri-natal sensitive periods permanently masculinises certain aspects of neural and behavioural development directly relating to reproduction. It is also widely cited that the organising effects of early testosterone exposure on brain anatomy and function may mediate a variety of non-reproductive behaviours and cognitive skills. Evidence here however remains equivocal. The following sections will discuss the reported and hypothesised organisational influences of testosterone on the brain and cognition in both animals and humans.

1.6 Organising effects of testosterone effects on the brain in animals

Empirical research in animals allows for the manipulation of testosterone exposure for research purposes, and thus a high level of experimental control into its effects. Given the obvious ethical restraints on comparable research in humans, experimental evidence into the effects of sex hormones in animals provides a valuable resource for the study of the impact of testosterone on neural anatomy and cognition.

There is now a large body of empirical evidence which has identified evidence for sex difference in aspects of neural anatomy in animals, a consideration of which is beyond the scope of this thesis. Of crucial importance to this thesis, there is strong evidence in animals for an organising role of testosterone on a number of neural sexual differences in the size of specific regions, the density of the neurons within such regions, and patterns of neural connectivity. For example, male animals display a larger volume of structure/number of neurons as compared to females in the sexually dimorphic nucleus of the preoptic area (SDN-POA) (Gorski, 1984; Gorski, Gordon, Shryne, & Southam, 1978; Gorski, Harlan, Jacobson, Shryne, & Southam, 1980; Jacobson, Csernus, Shryne, & Gorski, 1981), medial amygdala (Mizukami, Nishizuka, & Arai, 1983; Nishizuka & Arai, 1981), medial posterior subdivision of the bed nucleus of the stria terminalis (BSTPM) (Del Abril et al., Segovia, & Guillamon, 1987; Guillaman, Segovia, & Del Abril, 1988), accessory olfactory bulb (AOB) (Segovia, Orensanz, Valencia, & Guillamon, 1984; Valencia, Segovia, & Guillamon 1986), and the ventromedial nucleus of the hypothalamus (VMH) (Matsumoto & Arai, 1983, 1986). In all of these structures, research has demonstrated that neonatal gonadectomy decreases the volume and/or the number of neurons in males and an increase in the volume and/or the number of neurons is seen in females following neonatal androgenisation (Garcia-Falgueras et al., 2005).

Converse neural sexual dimorphisms, whereby females demonstrate a larger volume of structure/ number of neurons in comparison to males, have been reported in the; anteroventral periventricular nucleus (AVPV) (Arai et al., 1993; Bleier, Byne, & Siggelkow, 1982; Simerly, Swanson, Honda, & Gorski, 1985), medial anterior division of the bed nucleus of the stria terminalis (BSTMA) (Del Abril et al., 1987), the lateral anterior division of the bed nucleus of the stria terminalis (BSTLA) (Guillaman, Segovia, & Del Abril, 1988), parastrial nucleus (Del Abril et al., 1990) and the locus coeruleus (LC) (Guillamon, Del Blas, & Sergovia, 1988; Pinos et al., 2001; Rodriguez-Zafra et al, 1993). In these structures, male gonadectomy has been observed to increase the number of neurons, while female androgenisation produces the opposite effect. In rats there is also evidence for sexually dimorphic patterns of structural cortical asymmetry that are likely to be influenced by the actions of prenatal testosterone (PT). For example, in male rats the right hemisphere of the cerebral cortex is typically thicker than that of the left, while the reverse pattern of structural lateralization is generally observed in females (Diamond, Johnson, & Ingham, 1975). In female rats overiectomized at birth, a masculine pattern of cortical asymmetry is displayed (Diamond et al., 1981), while male rats castrated at birth display a female pattern of structural asymmetry (Diamond, Johnson, & Ehlert, 1979).

Notably, the majority of findings described above refer specifically to regions directly associated with reproduction and reproductive behaviours. Thus, while obvious connections can be made between exposure to testosterone and reproductive behaviours, the non-reproductive behavioural and cognitive consequences that may follow from such structural differences are not well understood. At present there is a relative paucity of research exploring the potential effects of testosterone on sex differences in areas of the brain not directly associated with reproductive behaviours, hence the current lack of evidence does not necessarily preclude a possible impact of testosterone in non-reproductive regions. It is also important to note that in addition to structural influences, comparative research also suggests that testosterone may influence sex differences in neurochemistry (see, De Vries & Simerly, 2002). In light of such findings it is important to recognise that the potential impact of testosterone on both reproductive and non-reproductive behaviours is likely to be even more widespread and complex than can be predicted in terms of ‘simple’ anatomical effects.

1.7 Organising effects of testosterone on cognition in animals

With regard to the potential cognitive effects of testosterone in animals, research suggests that male rodents consistently outperform females on a variety of spatial navigation tasks (e.g. Barrett & Ray, 1970; Beatty, 1979, Dawson, Cheung & Lau, 1975; Einon, 1980; Isgor & Senegelaub, 1998; Roof & Havens, 1992; Stewart, Skvarenina & Pottier, 1975). Evidence also implies that male and female rats use different navigational strategies in order to complete such tasks (Kanit et al., 1998; Williams, Barnett, & Meck, 1990; Williams & Meck, 1991; see Saucier, Bowman, & Elias, 2003 for evidence of similar effects in humans), although the extent to which differences in strategy may account for differences in performance remains an issue of debate. Research in rodents has revealed that neonatal castration in males and prenatal or neonatal androgen treatment in females can reverse typical sex differences in both performance and strategy on a variety of maze navigation tasks (Dawson et al., 1975; Isgor & Senegelaub, 1998; Joseph, Hess, & Birecree, 1978; Roof & Havens, 1992; Stewart, Skvarenina, & Pottier, 1975; Williams et al., 1990). Evidence has also been reported for a link between testosterone and cognitive performance in non-human primates. In male rhesus monkeys for example, castration or treatment with chemicals that block testosterone or its metabolites resulted in feminized performance on a visual discrimination learning task (Bachevalier & Hagger, 1991). Evidence also exists implicating the use of different strategies in male and female rhesus monkeys to solve spatial tasks and a potential role for PT in the use of such strategies (Herman & Wallen, 2007).

Taken together, comparative research certainly provides compelling evidence as to a potential effect of testosterone on certain aspects of cognition and neural anatomy. Crucially however it is important to acknowledge that the precise mechanism via which testosterone may be exerting its effects on such skills is an issue of debate. Certainly in rodents, with regard to both neural organisation and cognitive masculinisation, evidence suggests that the impact of testosterone may be critically dependent upon the aromatization of testosterone to estrogen, although growing research does suggest that the organisation of spatial ability may also be partially accomplished via the direct action of androgens (Jones & Watson, 2005). It is also vital to recognise that clear etiological comparisons are inherently complicated given evidence that the timetable of sensitive periods for neural development and brain maturation can show dramatic species dependant differences (Cohen-Bendahan et al., 2005).

1.8 Organising effects of testosterone on the brain in humans

In humans, sex differences have been revealed in the interstitial nucleus of the anterior hypothalamus (INAH 1-4) (an area of the brain thought to be homologous to the sexual dimorphic nuclei of the hypothalamus (SDN) in animals) and the bed nucleus of the stria terminalis (BNST) (Gorski, 2002; Swaab et al., 2002). Research suggests however that the morphologic sex difference in the INAH may not be established until the first postnatal years, and that the sex difference in the BNST appears only in adulthood (Swaab et al., 2002). Any potential impact of testosterone in these areas therefore appears not to be a consequence of prenatal effects alone.

Sex differences favouring males in overall brain size, as well as total grey and white matter volume are also widely reported (e.g. Allen**,** Damasio, Grabowski, Bruss, & Zhang, 2003; Filipek, Richelme, Kennedy, & Caviness, 1994; Nopoulos et al., 2000). When controlling for overall brain size there is evidence to suggest that females may demonstrate a greater proportion of grey matter as compared to males, and that males may demonstrate higher volumes of white matter as compared to females (Allen et al., 2003; Gur et al., 1999); sex differences in grey matter proportion however have not been consistently reported (Luders, Steinmetz, & Jancke, 2002).

In accordance with findings in rats, sex differences in aspects of hemispheric asymmetry have also been identified. According to Gur et al. (1999) males demonstrate a greater percentage of grey matter in the left hemisphere as compared to the right, a greater percentage of cerebral spinal fluid in the right hemisphere as compared to the left, and asymmetric patterns of white matter. No significant asymmetries were identified in females (Gur et al., 1999). Although controversial, there is also evidence for sex differences in the area, shape and fibre composition of the corpus callosum, the white matter fibre tract connecting the two cerebral hemispheres. As information between the two cerebral hemispheres is thought to be shared via the corpus callosum, this structure is thought to be associated with functional cortical lateralization and degree of interhemispheric connectivity. For example, females are typically reported to demonstrate a larger, more bulbous medial and posterior section (splenium) as compared to males (Allen**,** Richey, Chai, & Gorski, 1991; Davatzikos & Resnick, 1998; De Lacoste-Utamsing & Holloway, 1982; De Lacoste-Utamsing, Holloway, & Woodward, 1986; Dubb, Gur, Avants, & Gee, 2003), and in turn hypothesised to display greater levels of interhemipheric connectivity (Gur & Gur, 2004). In line with this hypothesis there is evidence for sex differences in functional hemispheric lateralization with females reported to display a lesser degree of functional cortical asymmetry between the two cerebral hemispheres as compared to males (see Wisniewski, 1998). In addition to identified differences in degree of functional asymmetry (how much information is shared between the two hemispheres), sex differences have also been documented with regard to the direction of functional asymmetry (which hemisphere processes the information most efficiently, fastest or accurately). In this respect, males are typically reported to be right hemisphere dominant, while the female pattern of dominance is characterized by the left hemisphere (Wisniewski, 1998). This claim however remains controversial, particularly in light of evidence for a lack of sex differences in functional lateralization on both language (Boles, 2005; Sommer, Aleman, Bouma, & Kahn, 2004; Sommer, Aleman, Somers, Boks, & Kahn, 2004, 2008) and visuospatial tasks (Kimura, 1999).

Sex differences in structural and functional cerebral lateralization are particularly pertinent to the discussion of the potential neural and cognitive impact of PT, as a number of theorists assert the possibility that foetal testosterone may act upon the brain during a critical period of development to influence hemispheric asymmetry (Geschwind & Galaburda 1987; Hines & Shipley 1984; Witelson, 1991). Geschwind and Galaburda (1987) controversially (see Bryden, McManus, & Bulman-Fleming,

1994) hypothesised an excess or reduction of *in utero* testosterone to slow the migration and maturation process of certain areas of the left hemisphere. Geschwind and Galaburda (1987) further postulated that such inhibited maturational development may result in compensatory growth in regions of the right hemisphere as well as some adjacent areas of the left hemisphere. Based on their theory, Geschwind and Galaburda (1987) implicated high PT levels in the aetiology of left-handedness, autism, dyslexia, migraine, stammering and links between cerebral lateralisation and disorders of the immune system resulting from effects on the thymus. An alternate theoretical stance for the relationship between cerebral lateralisation and testosterone comes from Witelson and colleagues (Witelson, 1991; Witelson & Nowakowski, 1991). Based largely on findings that in males only consistent right-handers have a smaller corpus callosum than non-right handers, they further hypothesised that testosterone in males may mediate axonal pruning ultimately resulting in greater lateralisation of cognitive function, with differential or more subtle effects in operation in females.

The potential effect of testosterone on sex differences in lateralization are often cited as one possible explanation for sex differences in cognitive processing in humans. At a general level of interpretation, verbal (in particular speech production), and fine motor functions are predominantly thought to be left hemisphere lateralised, whereas spatial abilities, complex visuo-spatial analysis, and certain aspects of emotional processing and prosodic interpretation, are typically right hemisphere lateralised (Hellige & Longstreth 1981; McGowan and Duka, 2000; Springer & Deutsch 1997; Teuber 1974). Based on such evidence, observed sex differences in the degree of hemispheric lateralization therefore (right hemispheric dominance in males and left hemispheric dominance in females) (Wisniewski, 1998) may account for an apparent male advantage in certain aspects of spatial task performance and an apparent female advantage in certain aspects of verbal task performance (Halpern, 2000). Importantly however, despite being widely cited it is also recognised that the majority of cognitive process are likely to show bilateral representation. Rarely can performance on any cognitive task be wholly attributed to one specific region or hemisphere. Furthermore, the particular brain areas associated with specific cognitive skills can vary across task and within task, depending on what particular aspect of a task is being completed.

In humans, the link between PT exposure and neural development cannot be experimentally established for obvious ethical reasons. With regards to cognition, various alternative indirect approaches have been utilised in order to examine the potential effects of testosterone on cognitive processing (discussed below).

Unfortunately, these methodologies have yet to be widely adopted in order to explore the possible actions of PT exposure on human brain structure and anatomy. Thus, while powerful manipulation of the prenatal hormone environment in laboratory animals suggests that the mechanism of neural sexual differentiation are orchestrated by gonadal steroids (e.g. Gorski, 2002; Swaab et al., 2002), the potential influence of PT on neural anatomy and possible sexual dimorphisms in neural structure in humans is not yet certain. At present a link between potential sex differences in structural neural anatomy and testosterone exposure in humans can only be speculated based on the notion that the effects of testosterone are similar across both human and animal species. Importantly however, it is vital to recognise that sex differences in the structure, function or activation of the brain, testosterone induced or otherwise, do not necessarily translate to sex differences in performance. Evidence actually suggests that males and females may engage different constellations of brain regions in order to achieve the same level of performance on at least some cognitive and intelligence measures (Haier, Jung, Yeo, Head, & Alkire, 2005).

1.9 Organisational impact of testosterone on cognition in humans

1.9.1 Evidence from clinical groups

One method that has been employed in order to explore the potential cognitive effects of PT exposure in humans has been to investigate the correlates of atypical hormone exposure as a result of prenatal hormone abnormality, so called “experiments of nature”. In this context, the most commonly studied disorder is congenital adrenal hyperplasia (CAH) (Pasterski et al, 2007).

CAH is an autosomal recessive disorder involving an enzymatic deficiency in the glucocorticoid pathway (White & Speiser, 2000) which (due to negative feedback) ultimately results in excessive exposure to androgens beginning prenatally (Miller & Levine, 1987). Because of their high level of androgen exposure, females with CAH are typically born with virilised external genitalia ranging in severity from ‘mild’ to ‘extensive’. Females with the condition however are typically assigned and reared as females, with their external genitalia surgically feminized early in development (Pang, 1997; Pasterski et al, 2007).

There is evidence that females diagnosed with CAH show male levels of performance on tests of spatial ability in both adolescence and young adulthood and in childhood (Hampson, Rovet, & Altmann, 1998; Puts, McDaniel, Jordan, & Breedlove, 2008; Resnick, Berenbaum, Gottesmen, & Bouchard, 1986). These findings however are not consistently replicated, with a number of studies reporting poorer performance in CAH females or no difference in performance in comparison to unaffected controls on spatial tests (Baker & Ehrhardt, 1974; Helleday, Bartfai, Ritzen, & Forsman, 1994; Hines et al., 2003; McGuire, Ryan, & Omenn, 1975). In direct opposition to findings with females, there is some implication that males with CAH may demonstrate poorer spatial abilities than controls (Hampson et al., 1998; Hines et al., 2003; Puts et al., 2008). Again however this finding is not consistently reported (Resnick et al., 1986). Where effects are revealed with regard to spatial ability, evidence that spatial skills may be enhanced in CAH females and reduced in CAH males may indicate a curvilinear relationship between testosterone exposure and spatial aptitude, whereby optimal levels of testosterone exposure for the expression of spatial ability reside in a low-to-normal male range. Studies exploring verbal functioning in females with CAH have reported no significant differences between CAH females and controls on measures of verbal fluency (Baker & Ehrhardt, 1974; McGuire et al., 1975; Resnick et al., 1986; Sinforiani et al., 1994). There is evidence however that females with CAH are more likely to be left-handed as compared to controls (Nass et al, 1987), although this difference is small and again, not consistently reported (Resnick, 1983).

One important limitation specific to the study of CAH relates to the fact that females with CAH invariably exhibit masculinised genitalia. It is possible therefore that any masculinisation of cognition and/or behaviour may purely be a consequence of social experiences and differential parental treatment resulting from their condition. While evidence generally suggests that parents do not exhibit differential treatment towards their daughters with CAH as compared to their unaffected sisters (Berenbaum & Hines, 1992; Ehrhardt & Baker, 1974) these findings require further confirmation. At present, the possibility that differences in social responses may account for any apparent masculinisation of behaviour and cognition cannot be eliminated.

A second clinical condition that offers a ‘natural experiment’ into the potential behavioural and cognitive effects of androgens is the rare endocrine disorder Androgen Insensitivity Syndrome (AIS). Individuals with AIS demonstrate the typical 46,XY male karyotype and thus the SRY gene initiates male sexual differentiation including development of the testes (Cohen-Bendahan et al., 2005; Danilovic et al., 2007).

Despite being exposed to normal levels of testosterone a lack of functioning androgen receptors due to a X linked defect on the androgen receptor gene results in an insensitivity to the actions of all androgens both pre- and postnatally, and in turn, varying degrees of defective masculinization. The degree of androgen insensitivity can range from complete to partial. In the case of complete AIS (CAIS) individuals exhibit female-typical external genitalia and are consequently assigned and raised as females (Collaer & Hines, 1995).

Evidence suggests that individuals with CAIS demonstrate poorer performance on a number of visuospatial tests (Imperato-McGinley, Pichardo, Gautier, Voyer, & Bryden, 1991). While individuals with the condition are invariably raised as females and thus, the possibility exists that differences between CAIS and control males may arise purely a result of female typical-socialization, this suggestion is highly unlikely given evidence that; *a)* CAIS males demonstrate lower spatial task performance relative to both male and female controls and *b)* differences in performance revealed between CAIS males and control males are not always mirrored by similar differences between control males and females (Imperato-McGinley et al., 1991). As noted above, typically developing females demonstrate androgen receptivity and are exposed to low levels of circulating androgens both pre- and postnatally, as a consequence males with CAIS actually experience lower levels of androgen exposure than 'typical' females. The pattern of results described above therefore suggest a progressive increase in spatial skill with increasing androgen exposure, i.e. from entirely absent exposure (CAIS males) to low exposure in a female typical range (control females) to moderate exposure in a 'normal' male range (males controls). This is line with revealed effects from CAH where research suggests improved spatial task performance in females exposed to excessive androgens.

Finally, the cognitive and behavioural profiles of individuals with the condition Idiopathic Hypogonadotropic Hypogonadism (IHH) have also been examined in an attempt to explore the potential impact of testosterone on development. IHH is characterized by low plasma testosterone levels secondary to a deficiency in hypothalamic gonadotropin-releasing hormone (GnRH) (Hampson, 1995). While the condition is evident in both males and females the majority of research focuses exclusively on potential behavioural and cognitive correlates in males. Males with the condition demonstrate the normal 46XY karyotype and, despite severe androgen deficiency, display typical male external genitalia and are raised as boys. IHH can be both congenital or develop later in life (Collaer & Hines, 1995). In the case of delayed development however the condition typically does not present until individuals reach their late 20s to early 30s following normal pubertal development (Whitcomb & Crowley, 1993).

There is evidence that IHH males show visuospatial deficits. Impaired performance has been reported on the Wechsler Block Design subtest, the Embedded Figures Test, the Space Relations subtest of the Differential Aptitude Test and the Rod and Frame Test (Buchsbaum & Henkin, 1980; Hier & Crowley, 1982). Research also suggests that IHH males may show poorer spatial memory compared to controls (Buchsbaum & Henkin, 1980; Kertzman, Robinson, Sherins, Schwankhaus, & McClurkin, 1990). Importantly however, IHH deficits on the Wechsler Block Design subtest and alternative tests of visuospatial ability are not consistently replicated (Buchsbaum & Henkin, 1980; Cappa et al., 1988). Where effects are evident, findings are generally in agreement with those discussed above regarding individuals with CAH and CAIS.

Evidence from all three clinical groups implies impaired spatial task performance where testosterone exposure is below ‘typical’ male levels, as compared to those individuals where testosterone exposure is in a high female/’normal’ male range. Also in agreement with previous research there is evidence that the severity of androgen deficiency in IHH males correlates positively with the degree of visuospatial deficit (Hier & Crowley, 1982). Hier and Crowley (1982) further report an absence of visuospatial deficits in late onset hypogonadaism following normal puberty, suggesting a critical importance of early hormone exposure in the development of visuospatial skills. Research that has been conducted to date has failed to find differences between IHH males and controls on measure of verbal ability (Cappa et al., 1988; Hier and Crowley, 1982) and hemispheric lateralization for linguistic processing (Cappa et al., 1988), although one study does report a deficiency in IHH males on a measure of verbal fluency (Cappa et al., 1988). There is however evidence that IHH males may show deficits as compared to controls on tasks that measure aspects of visual and verbal memory (Cappa et al., 1988).

There are a number of important limitations relevant to the study of any clinical condition which should be considered when evaluating the findings described above. Firstly, sample sizes are invariably low; many studies in clinical populations therefore have low power and thus an increased risk of Type II error. Secondly, individuals with an endocrine disorder may have further medical issues and/or additional hormonal imbalances beyond prenatal androgens, which may at least partially account for differences with controls. In the case of CAH for example, individuals with the condition also demonstrate abnormal progesterone and corticosteroid levels (Cohen-Bendahan et al., 2005) which may also contribute to atypical maturation of the central nervous system. In addition to issues surrounding the presence of alternative hormone imbalances, the study of clinical conditions in the context of the potential behavioural/cognitive impact of testosterone is complicated further due to the fact that all endocrine disorders studied are a result of abnormal androgen exposure in general. It is impossible therefore to identify the relative contribution of testosterone vs. other androgens such as dihydrotestosterone, a metabolite of testosterone in any revealed effects. Finally, inevitable problems exist with regard to generalisation. While any revealed effects may implicate an impact of excessive testosterone exposure in one specific subgroup they actually provide little information about the effects of typical variation.

1.9.2 Evidence from umbilical cord blood and amniotic fluid

In addition to research from clinical populations where prenatal hormone exposure is atypical for a person's sex, there is also evidence into the early effects of testosterone on the brain and cognition in normal populations from direct measures of peripheral testosterone obtained via umbilical cord blood and amniotic fluid.

Jacklin, Wilcox, & Maccoby, (1988) explored the potential relationship between concentrations of five steroid hormones (testosterone, androstenedione, estradiol, estrone and progesterone) assayed via umbilical cord blood, and performance on four cognitive subtests (reading, numerical skill, listening and spatial ability) at 6 years of age. Findings revealed a significant negative relationship between umbilical cord blood androgen levels (testosterone and androstenedione) and spatial ability in girls. In contrast to the implications of previous evidence in clinical populations, these results suggest an association between higher prenatal androgen exposure and lower spatial skill in females. The findings revealed no significant correlations in males.

There are issues surrounding the use of umbilical cord blood which suggest that the measure may not be an entirely reliable reflection of PT exposure. Firstly, the sensitive period for the behavioural effects of hormones is believed to occur during approximately weeks 8 – 24 of gestation (Cohen-Bendahan et al., 2005; Collaer & Hines, 1995). Hormones from the umbilical cord therefore are unlikely to represent levels of fetal testosterone exposure during hypothesised critical periods of brain maturation. More recent research actually reports no significant correlation between testosterone concentration assessed in umbilical cord serum at birth and amniotic fluid samples during weeks 15-18 of pregnancy (van de Beek, Thijssen, Cohen-Kettenis, van Goozen, & Buitelaar, 2004). In addition, the onset and stress of the labour and delivery process itself may actually affect hormone levels (Fuchs & Fuchs, 1984) thus distorting the extent to which hormone concentrations in umbilical cord blood reflect levels experienced throughout pregnancy. Secondly, while sex differences in umbilical blood hormone levels have been identified (Dawood & Saxena, 1977; Jacklin et al., 1988; Sakai, Baker, Jacklin, & Shulman, 1992) they are small and not consistently detected (Maccoby, Doering, Jacklin, & Kraemer, 1979; van de Beek et al., 2004). Given that the critical period for the behavioural effects of hormones is thought to coincide with the early-mid gestational peak in male testosterone, sex differences during this period would be expected to be fairly large. A lack of sex differences in levels of testosterone exposure assessed via umbilical cord blood therefore once again calls into question the extent to which hormone concentrations in the umbilical cord reflect levels of testosterone exposure during the hypothesised prenatal sensitive period of neural development.

Studies investigating PT levels in amniotic fluid derived via routine amniocentesis may be more promising. Coinciding with peak serum fetal testosterone levels in males, amniotic fluid levels are typically obtained during what's thought to be the important developmental period of prenatal hormone effects i.e. the second trimester (usually approx. 14-20 weeks of gestation). Sex differences in amniotic fluid are consistently reported (Dawood & Saxena, 1977; Finegan, Bartleman, & Wong, 1989; Judd, Robinson, Young, & Jones, 1976; Lutchmaya, Baron-Cohen, & Raggatt, 2002; van de Beek et al., 2004), with the maximal sex difference in amniotic testosterone thought to occur during weeks 12-18 of gestation (Finegan et al., 1989). A study by Finegan, Niccols, & Sitarenios, (1992) examined relations between prenatal hormone levels from amniotic fluid during the second trimester (collected during routine amniocentesis between weeks 14 and 20 of gestation) and subsequent cognitive ability at 4 years of age. Contrary to the findings of research in clinical samples, but in line with the findings of Jacklin et al. (1988) discussed above, girls with low amniotic testosterone levels demonstrated higher average block-building scores as compared to those with higher amniotic testosterone levels. No significant association was revealed for boys. The authors noted however that the children may have been too young to permit reliable assessment of spatial abilities with the tools available at the time.

Grimshaw, Sitarenios, and Finegan (1995) reported a positive association between amniotic levels of testosterone and mental rotation task (MRT) performance in girls aged 7 years old. Broadly in line with evidence from clinical populations, girls with higher amniotic levels of testosterone had faster (but not necessarily more accurate) mental rotation performance than did girls with lower levels. A non-significant trend in the opposite direction was identified in boys. Importantly however, significant findings were only revealed in a small subgroup of 12 girls who specifically used a rotation strategy (characterized by relationship between reaction time and figure orientation) and, contrary to the suggestion of previous research regarding sex differences in adult MRT performance (Linn & Petersen, 1985; Voyer et al., 1995), girls classified as using a rotational strategy were actually faster at rotation than boys who had also adopted a rotation strategy. Findings from Finegan et al. (1992) also demonstrated PT in girls to show a curvilinear relationship (inverted U-shaped) between language comprehension and classification abilities.

Grimshaw, Bryden, and Finegan (1995) explored the potential relationship between indicators of lateralization (measured by handedness and by dichotic listening task performance) to amniotic fluid testosterone levels in children aged 10 years old. For girls, amniotic fluid testosterone was positively correlated with degree of right-handedness and degree of left-hemisphere lateralization (right-ear advantage) for language; thus higher levels of testosterone were associated with stronger right-handedness and stronger left-hemisphere language representation. For boys, testosterone was positively correlated with degree of right-hemisphere specialization (left-ear advantage) for the recognition of emotion. Results in both sexes were interpreted to be most consistent with the hypothesis of Witelson and Nowakowski (1991) that high levels of testosterone exposure *in utero* leads to greater lateralization of function. A study by Lutchmaya et al. (2002) reported a significant negative relationship between amniotic fluid levels of fetal testosterone and vocabulary size in infants aged 18-24 months when data from both sexes was examined together. No significant relationships however were revealed for analysis within-sex.

As with all methods there are limitations to the use of amniocentesis samples. Firstly, problems exist with regard to generalizeability. As amniocentesis is normally conducted for the purpose of diagnosing fetal abnormalities samples are opportunistic. Women referred for amniocentesis tend to be upwardly skewed with reference to age and educational level (van de Beek et al., 2004). Also as amniotic fluid studies rely on selective populations they tend to have small sample sizes. An alternative problem stems from the fact that the origins of androgens in amniocentesis are not yet fully understood. While it is known that hormones appear to enter amniotic fluid primarily via diffusion through the fetal skin in early pregnancy and via fetal urine in later pregnancy (Judd et al., 1976; Schindler, 1982) there is evidence to suggest that the relationship between amniotic testosterone levels and peripheral blood levels may be low (Rodeck, Gill, Rosenberg, & Collins, 1985). While sex differences in amniotic testosterone levels are consistently identified, such research also suggests that they may be smaller than those revealed in peripheral blood, as measured in abortuses (Rodeck et al., 1985). Further research is needed to fully elucidate the extent to which amniotic fluid levels of testosterone truly represent fetal exposure. Finally, a general methodological issue applicable to studies utilising measures from both umbilical cord blood and amniotic fluid is that such research require years of investment before cognitive measures can be obtained and meaningful results are available. As a consequence of their prospective nature both amniotic fluid and umbilical cord blood studies often suffer from a high drop-out rate and the associated problem of selective attrition (Cohen-Bendahan et al., 2005).

1.9.3 Evidence from opposite sex twins

There is good evidence in animals that an individual's position in the uterus with regard to the sex of its littermates (intrauterine position - IUP) can influence subsequent behaviour and physiology. Female animals that develop between male fetuses *in utero* appear masculinised in terms of postnatal behaviour, anatomy, and reproductive characteristics as compared to those that develop between female fetuses (Clark & Galef, 1998; Rohde Parfet et al., 1990; Ryan & Vandenbergh, 2002; vom Saal, 1989). Similarly there is evidence that male animals that develop between two females show less masculine reproductive characteristics and sexual behaviours, and lower levels of aggression than those that developed between two males (Beatty, 1992; Clark & Galef, 1998). As the studies of IUP effects are consistent with studies in which hormones are manipulated directly and shown to effect later behaviour, the masculinising effect in females is attributed to the transfer of testosterone from the male fetus to the female fetus (Even, Dhar, & vom Saal 1992).

In parallel to the IUP effect in animals it has been suggested that human twins may also be affected by the sex of the co-twin (Miller, 1994; Resnick, Gottesman, & McGue, 1993). As a result of sharing the womb with a male co-twin female members of opposite twin pairs are assumed to be exposed to higher levels of PT as compared to female members of same sex twin pairs. Similarly, male members of opposite-sex twin pairs are assumed to be exposed to lower lever of PT as compared to same-sex twin pairs. Thus the study of cognitive correlates in opposite-sex twins has been highlighted as one alternative method for exploring the organising impact of testosterone.

For example, there is evidence that females with male co-twins demonstrate better spatial ability on the Mental Rotations Test than same-sex twins (Cole-Harding, Morstad, & Wilson, 1988). This finding is in line with evidence from clinical populations suggesting that heightened exposure to testosterone may have a facilitative effect on spatial skill in females. No significant sex differences have been revealed between opposite-sex and same-sex twin pairs in terms of handedness (Elkadi, Nicholls, & Clode, 1999). Cohen-Bendahan, Buitelaar, van Goozen, and Cohen-Kettenis (2004) did however report a significant difference in cerebral lateralization between 10-year old twins, such that opposite-sex twin girls displayed a more masculine pattern of lateralization than same sex twin girls, reflected in a larger right ear advantage. These results again suggest that testosterone may increase lateralization of function. Importantly however these results were not replicated in a re-evaluation of the findings in the same sample of girls at 13 years of age (Cohen-Bendahan, 2005).

Once again methodological problems exist. Firstly, and perhaps most obviously, twins share postnatal environments. As it is likely that opposite-sex twin pairs may be exposed to different gender-related social environments than same-sex pairs, it is almost impossible to tease apart the relative contribution of potential hormonal and social-experiential influences on behaviour and cognition. Secondly, in humans, the level of hormonal transfer between twins remains unknown. As identified above, the hypothesis for a potential effect of testosterone in human opposite-sex twins is based largely on evidence from IUP effects in animals. Importantly however, animal litters are invariably larger than two. As a result, evidence from animals is typically based on male-female-male IUP patterns, and thus likely to demonstrate much stronger effects than those observed in humans. Finally, the precise mechanisms of hormonal transfer in multiple births and changes in the possible mechanisms of potential hormonal transfer over fetal development (e.g. the potential effects of changes in fetal skin from permeable in early pregnancy to non-permeable in later pregnancy) remain poorly understood.

1.9.4 Evidence from somatic markers

The final alternative method for investigating the potential behavioural and cognitive effects of prenatal exposure to testosterone involves the use of morphological indices as potential proxies for *in utero* testosterone exposure. One such morphological indicator is otoacoustic emissions (OEAs). OAEs are weak sounds emanating from the inner ear (cochlear) that can be recorded using a miniature microphone system inserted into to the external ear canal (Cohen-Bendahan et al., 2005; Loehlin & McFadden, 2003; McFadden, Loehlin, & Pasanen, 1996). They are thought to be stable from an early age (Burns, Campbell, Arehart, & Keefe, 1993; Burns, Campbell, & Arehart, 1994; McFadden, 2002) and show reliable sex differences with females demonstrating both louder and more frequent OAEs as compared to males (Bilger et al., 1990; Loehlin & McFadden, 2003; McFadden & Champlin, 2000; McFadden & Pasanen, 1998, 1999; Talmadge, Long, Murphy, & Tubis, 1993). Based on revealed sex differences in OAEs and evidence that opposite sex dizygotic twins demonstrate masculinised OEA patterns relative to controls (McFadden & Loehlin, 1995; McFadden et al., 1996) it has been suggested that patterns of OAE expression may be associated with level of exposure to fetal androgens (McFadden, 1993; McFadden, 1998; McFadden, 2002; McFadden et al., 1996).

To the author’s knowledge, only one study to date has been conducted exploring a potential relationship between OAEs and sex-typed cognitive performance thought to be influenced by pre-natal testosterone. Loehlin and McFadden (2003) reported no significant relationship between OEAs and scores on the Mental Rotation Task (MRT), angle of water surface estimation task, self reported sense of direction, and self reported mechanical ability. The same study reported no significant relationship between OAEs and a measure of laterality (a measure of combined handedness and footedness).

Currently, research utilising OAEs as a potential marker of prenatal androgen remains fairly limited. The majority of evidence has focused on their possible association with sexual orientation, and all of the studies conducted so far appear to have been carried out by the same research group. With regard to methodology, assessment of OAE expression requires 20 minutes of recording in a quiet environment. This can present difficulties when working with young children, particularly as body movements can interfere with data collection during recording (Cohen-Bendahan et al., 2005). There is also evidence that OAE expression may be influenced by menstrual cycle fluctuations (McFadden, 1998) and oral contraceptive use in females (McFadden, 2000; McFadden, 2002). It is difficult to determine therefore whether any revealed effects are associated with prenatal or adult sex hormone fluctuations. Ultimately, more research is needed in order to evaluate the usefulness of OAEs as a reflection of androgen exposure over the fetal period.

Another morphological measure thought to reflect prenatal exposure to sex hormones is the pattern of dermal ridges that constitutes the human fingerprint (dermatoglyphics). Dermal ridges are fixed by the fourth month of gestation (Holt, 1968) and show sex differences in terms of both finger ridge count (males have more total ridges than females) (Holt, 1968; Kimura & Carson, 1995) and asymmetry (although both sexes have more ridges on the right hand than on the left hand - R>L, there is evidence that the reverse asymmetry of left greater than right - L>R is more common in females than in males) (Kimura & Carson, 1995; Sanders & Kadam, 2001). Studies show that individuals with a R>L pattern may perform better on some tasks on which males generally excel, while that individuals demonstrating a L>R pattern may show enhanced performance on tasks which generally favour women (Kimura & Carson, 1995; Kimura & Clarke, 2001; Sanders & Waters, 2001). Similar findings have been replicated in children (Sanders & Kadam, 2001).

At present however there is little evidence directly associating dermatoglyphics with levels of PT exposure. As total ridge count is inversely associated with the amount of material in the sex chromosomes (such that individuals with only one X chromosome, as is the case in some individuals with Turners syndrome, have the highest count of all) sex differences in the trait may be entirely unrelated to sex hormones. Also, with regard to sex differences in patterns of asymmetry, evidence for a female preponderance in L>R patterns of dermatoglyphics is not consistently reported (Holt, 1968; Slabbekoorn, van Goozen, Sanders, Gooren, & Cohen-Kettenis, 2000).

Perhaps the most well know somatic marker associated with PT is handedness. The majority of people show a right hand preference for writing and other skilled manual tasks, this preference however is more prevalent in women than men (Calnan & Richardson, 1976; Lansky et al., 1988, Oldfield, 1971). An individual’s preference for the use of their right or left hand has long been viewed and an indication of cerebral lateralization. While the majority of both left and right handers show left hemispheric dominance for language for example, evidence suggests that left handed individuals have a higher incidence of atypical (right hemispheric or mixed) language representation (Knecht et al., 2000; Pujol, Deus, Losilla, & Capdevila, 1999). As described previously, the work of Geschwind and Galaburda (1987) hypothesised that high levels of testosterone exposure prenatally may damage or slow development of the typically dominant left hemisphere resulting in compensatory growth in areas of the right hemisphere. Geschwind and Galaburda (1987) further postulated that this mechanism causes a shift in certain left-hemispheric functions to the right resulting in decreased lateralization and increased left handedness with increasing exposure to fetal testosterone. In support of this hypothesis there is some evidence for increased rates of left handedness in girls with CAH (Nass et al., 1987, see above) and for a relationship between one proposed marker of PT (the second to fourth digit ratio, 2D:4D – see subsequent section) and increased rates of left handedness or decreased degree of right handedness (Fink, Manning, Neave, & Tan, 2004; Manning, Trivers, Thornhill, & Singh, 2000; Manning & Peters, 2009; Nicholls, Orr, Yates, & Loftus, 2008). Contradictory findings have however been reported by Grimshaw, Bryden, & Finegan, (1995) who observed a significant positive correlation between degree of right handedness and level of amniotic fluid testosterone in girls (see section 1.9.2).

Based on the assumption that handedness may reflect certain aspects of brain organisation a number of authors have attempted to relate hand preference to systematic differences in cognitive skill. However, evidence regarding this relationship remains equivocal. Although some studies demonstrate a general advantage for right handed-individuals (e.g. Levy, 1969; McManus & Mascie-Taylor, 1983) others report no such differences (Nettle, 2003; Newcombe et al., 1975). There is also research to suggest that ambidextrals or individuals possessing a weak hand preference may perform more poorly than consistent left or right handers on certain cognitive tasks (Crow, Crow, Done, & Leask, 1998; Peters, Reimers, & Manning, 2006; but see Mayringer & Wimmer, 2002).

An increased proportion of left handedness has been associated with certain developmental disorders such as autism (e.g. Dane & Balci, 2007; McManus, Murray, Doyle, & Baron-Cohen, 1992) and dyslexia (e.g. Eglinton and Annett, 1994), the link between handedness and dyslexia however is not consistently reported (e.g. Bishop, 1990). Intriguingly an excess of left handedness has also been associated with a precocious ability in certain cognitive domains such music (Aggleton, Kentridge**,** & Good, 1994; Hassler & Gupta, 1993), mathematics (Benbow, 1986, 1988, see section 1.10) and art (Noroozian, Lotfi, Gassemzadeh, Emami, & Mehrabi, 2002).

Crucially however, the precise underlying mechanisms responsible for revealed hand preferences remain a source of debate. As well as a possible influence of PT, the trait is also believed to be affected by maternal handedness, family history of sinistrality (Annett, 1998; 1999) and history of early brain damage (Rasmussen & Milner, 1977). As a result, the underlying mechanisms of any potential relationship between handedness and cognitive function cannot be clearly interpreted in the context of a possible influence of PT. Unsurprisingly therefore, handedness is rarely adopted as a potential marker of PT in order to assess potential relationships between exposure to the hormone and subsequent cognitive functioning.

1.9.5 Second to fourth digit ratio (2D:4D)

The ratio between the length of the 2nd (index) and 4th (ring) fingers (2D:4D, digit ratio) is a sexually dimorphic trait. In most populations males demonstrate a significantly lower mean 2D:4D ratio, and therefore tend to have a longer fourth finger compared to their second; in females this pattern is typically reversed (George, 1930; Manning, Barley, et al., 2000; Phelps, 1952). 2D:4D is thought to be determined *in utero* by around the 14th week of gestation, thus coinciding with the hypothesised period of PT effects, (Garn, Burdi, & Babler, 1975; Manning, Scutt, Wilson, & Lewis-Jones, 1998; Phelps, 1952) and is sexually dimorphic from an early age (McIntyre, Cohn, & Ellison, 2006; Trivers, Manning, & Jacobson, 2006). Sex differences have also been reported in the second to fifth finger and the third to fourth finger, both in the same direction as the 2D:4D ratio, i.e. with males demonstrating a lower ratio as compared to females (McFadden & Shubel, 2002). 2D:4D however, remains by far the most extensively studied of all digit ratios.

While sexual dimorphism in 2D:4D has been known for more than 125 years, it was only in 1998 that Manning and colleagues proposed that 2D:4D may act as a proxy marker for levels of testosterone (T) and estrogen (E) to which a developing fetus is exposed (Manning et al., 1998). Consistent with this hypothesis, Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning (2004) found that the ratio of prenatal T to E assessed via routine amniocentesis was negatively related to 2D:4D at age 2. Further support can be gleaned from evidence that; *i)* females diagnosed with Congenital Adrenal Hyperplasia (CAH) exhibit lower 2D:4D than unaffected controls (Brown, Hines, Fane, & Breedlove, 2002; Ökten, Kalyoncu, & Yaris, 2002; but see Buck, Williams, Hughes, & Acerini, 2003), *ii)* 2D:4D is positively correlated with androgen insensitivity (as measured by the number of CAG repeats in the androgen receptor gene) in males (Manning, Bundred, Newton, & Flanagan, 2003), *iii)* female opposite sex-twins, thought to be exposed to heightened androgen exposure from their male twin, show lower 2D:4D as compared to controls (van Anders, Vernon, & Wilbur, 2006), *iv)* an alternative positive correlate of testosterone in the mother i.e. the waist/hip ratio, is negatively correlated with the 2D:4D of their child (Manning, Trivers, Singh, & Thornhill, 1999), *v)* the 2D:4D of mothers is similar to that of their children and mothers in possession of low 2D:4D display high testosterone levels in the amniotic fluid of their foetuses (Manning, 2002).

Further support for a potential association between 2D:4D and PT exposure comes from evidence demonstrating a common genetic link between the formation of the digits and the genitals. The development of both the urogenital system including the gonads and the appendicular skeleton are under the control of the *Homeobox* or *Hox* genes (Zakany & Duboule, 1999). More specifically, evidence suggests that the posterior-most *Hoxd* and *Hoxa* genes similarly control the development of distal limbs i.e. finger and toe length and genital eminence (Williams, Greenhalgh, & Manning, 2003). Kondo, Zakany, Innis, & Duboule (1997) demonstrated in mice that the progressive removal of the posterior *Hox* gene function to result in concomitant loss of the digit and genital bud derivatives. Furthermore, *hox* gene mutations have been identified in expressions of the hand and foot genital syndrome involving several anomalies on distal limbs and genital buds (Manning & Bundred, 2000). As the development of genital structures is influenced by androgens, it is assumed that the genes coding for genital development must be directly or indirectly modulated by androgens. This in turn lends itself to the logical conjecture that development of other structures influenced by these genes, namely the digits, might also be modulated by androgen. Despite some inconsistencies in revealed relationships (see Putz, Gaulin, Sporter, & McBurney, 2004 for a review) numerous correlational studies have found significant associations between 2D:4D and a wide range of cognitive and behavioural factors, thought to be mediated by the actions of PT.

With regard to spatial ability, a number of studies have reported significant negative relationships between 2D:4D and mental rotation whereby lower ratios (higher PT) were associated with better task performance in males (Manning & Taylor, 2001; McFadden & Schubel, 2003; Peters, Manning, & Reimers, 2007; Sanders, Bereczkei, Csatho, & Manning, 2005). A web based study involving a very large sample conducted by Collaer, Reimers, & Manning (2007) also reported a significant negative correlation between 2D:4D and visuospatial task performance on a Judgement of Line Angle and Position (JLAP) task. These findings however are not consistently indentified, with a number of studies reporting no significant effects in males or associations in the opposite direction for spatial task performance (Austin, Manning, McInroy, & Mathews, 2002; Burton, Henninger, Hafetz, 2005; Coolican & Peters, 2005; Hampson, Ellis, Tenk, 2008; Kempel et al., 2005; Poulin, O'Connell, & Freeman, 2004; Putz et al., 2004). Findings in females are equally mixed. While the majority of research typically reports no significant association between 2D:4D and spatial task performance in women (Austin et al., 2002; Coolican & Peters, 2003; Hampson et al., 2008; Manning & Taylor, 2001; McFadden & Shubel, 2003; Poulin et al., 2004; Sanders et al., 2005; van Anders & Hampson, 2005) a number of authors have reported a significant negative relationship between 2D:4D and spatial competence whereby less female typical (lower) 2D:4D measures are associated with better task performance (Burton et al., 2005; Collaer et al., 2007; Csathó et al., 2001; Kempel et al. 2005; Peters et al., 2007). While, in line with the majority of previous evidence, Poulin et al. (2004) reported no significant relationship between 2D:4D and mental rotation task performance in females, the authors also found a significant positive correlation between right hand 2D:4D and scores on picture free recall and picture placement tasks, suggesting lower PT exposure to correspond with better performance on such tasks. Importantly however, while studies exploring a potential impact of PT on spatial ability typically employ male favouring spatial assessments both picture recall and picture placement task showed a significant female advantage. Only one other study reports an association between higher (more feminine) 2D:4D ratios and better spatial task performance (Putz et al., 2004).

In terms of behavioural asymmetries, research has shown low digit ratios (higher PT) in children to relate to left hand preference in a peg moving task in both boys and girls (Manning, Trivers, et al., 2000). Similar results are reported by Fink et al. (2004), who revealed significant associations between lower 2D:4D and reduced degree of right handedness in male and female children. In adults, a study by Nicholls et al. (2008) also reported lower 2D:4D ratios (higher PT) in non-dextral participants. Similar results have been reported by Manning and Peters (2009) in a recent internet study. These authors found a significant relationship between right hand 2D:4D and writing preference with low 2D:4D (higher PT) associated with left hand preference in a sample of over 170,000 participants. Contradictory findings however had been reported by

Ypsilanti, Ganou, Koidou, & Grouios (2008) who described evidence for low 2D:4D scores in right handed participants as compared to left handed participants, implying higher PT exposure in right handed individuals. With regard to functional cognitive lateralization a recent study by Bourne and Gray (2009) reports an association between lower 2D:4D and stronger right hemisphere dominance on two versions of a chimeric faces test and the landmark test whereby participant were ask to identify the longer side of a bisected line.

Both Austin et al. (2002) and Kempel et al. (2005) reported no significant association between 2D:4D and verbal ability in either males or females. Evidence from Luxen and Buunk (2005) however demonstrated a significant positive correlation between right hand 2D:4D and verbal IQ in male and female data combined and female data analysed separately but not male data analysed separately. Contradictory evidence is also presented by Burton et al. (2005) who reported verbal fluency and scores on the verbal American Scholastic Achievement Test in males to be associated with less of a male-typical left hand 2D:4D ratio (higher ratios – lower T), but no significant correlations between verbal ability and 2D:4D in females. In opposition to the findings of Burton et al. (2005), Brosnan (2008) found a significant positive correlation between 2D:4D and UK Standardised Assessment Task (SAT) literacy scores in females, this research however failed to identify the same relationship in males.

The use of 2D:4D as a tool for exploring the relationship between PT and its behavioural and cognitive correlates has been criticised on the basis that results are equivocal and at times contradictory. A study by Putz et al. (2004), for example, questioned the use of 2D:4D to investigate the potential behavioural and cognitive impact of PT. In an attempt to re-investigate and clarify the potential relationship between 2D:4D and behavioural and cognitive measures Putz et al., (2004) tested 57 correlations between 2D:4D and several variables thought to be influenced by prenatal sex hormones including, sexual orientation, spatial ability, status, physical prowess, components of reproductive success, voice pitch, sociosexuality, mating success and fluctuating asymmetry. While findings did reveal significant negative correlations between 2D:4D and sexual orientation in both males and females, and a significant positive correlation between 2D:4D and MRT task performance in females, 2D:4D was found to be unrelated to the majority of traits assessed. Putz et al. (2004) proposed that 2D:4D may be an index of PT exposure, but that PT has different effects on the two genders, and on different traits or individuals, due to the fact that androgens fluctuate during development, and dimorphic traits may differentiate at different times.

It is important to highlight however that if indeed Putz's conclusion is correct this issue is equally applicable to studies utilising measures of amniocentesis, umbilical cord blood, and alternative somatic markers to investigate a potential effects of PT on behaviour and cognition. It is also important to note that many of the inconsistencies in previous findings identified by Putz et al. (2004) could be related to differences in methodology across studies and small samples sizes. Given the extent of research that has found significant correlations between 2D:4D and traits associated with testosterone in the expected direction it would seem, at this point, to be premature to dismiss the use of 2D:4D as a potential biomarker for the possible behavioural and cognitive effects of testosterone on the basis of this study.

Crucially, the use of 2D:4D as a potential maker of PT has a number of important advantages over alternative measures. Firstly, it provides a simple and widely available method for examining hormonal effects on human behaviour. In particular the 2D:4D measurement can be used easily with children and measured/re-measured with ease and reliability by taking a permanent photocopy or scan of the hands. Secondly, evidence suggests that the ratio is unaffected by the changing levels of sex hormones throughout puberty (Manning et al., 1998). Any revealed effects therefore can be linked to early organising influences in a straightforward manner. Finally, in 'normal' populations, the method facilitates the use of large, controlled, representative samples, of any age.

1.10 Testosterone and mathematical ability

In evaluating the potential effects of PT on cognition, the previous section, focused specifically on aspects of verbal and spatial ability and functional and behavioural lateralization. Yet to be considered however is the widely hypothesised (e.g. Benbow & Stanley, 1983; Kimura, 1996; Kolata, 1983), yet far more scarcely investigated, link between PT exposure and mathematical ability.

Similar to verbal and spatial proficiency, sex differences in mathematical performance are widely identified. In the general population males have been reported to demonstrate accelerated mathematical learning (Leahey & Guo, 2001) compared to females, as well as superior performance on assessments of mathematical reasoning, analytical spatial-visualization, geometry, and statistics (Friedman, 1989; Jensen, 1988; Maccoby & Jacklin, 1974; Stones, Beckmann, & Stephens, 1982). Conversely, females have been reported to outperform males on tests of mathematical sentences, algebra and computation (Stones et al., 1982). The general consensus suggests however that any apparent sex differences in mathematical performance typically do not emerge reliably until approximately 13-16 years of age (Hyde et al., 1990). It is also generally cited that where sex differences do exist in mathematical tasks, a stronger male advantage is apparent with increasing cognitive level (Fennema, 1974).

In an educational context evidence from the Third International Mathematics and Science Study (TIMSS; Mullis et al., 1999) suggests that where sex differences in mathematics performance are revealed they predominantly reflect a male advantage. A male advantage in mathematics performance is reported to be most pronounced in advanced mathematics courses (Mullis et al., 1999). More recent data from the USA suggests that, while a sex differences in mathematics favouring boys appears to be reducing, a male advantage is still evident in grade-8 children (National Centre for Education and Statistics, 2004). In contrast, prior literature on the subject also from the USA, argues for a female advantage in mathematic performance at all levels (Kimball, 1989; Willingham & Cole, 1997). There is some evidence however that females may score lower when the content is not directly related to what is taught in the curriculum (see Geary, 1996; Halpern, 2000). Data from the UK demonstrates a female advantage in both GCSE and Advanced level examination performance in mathematics (Department of Education and Skills, 2002). It is important to recognise however that performance in educational context is likely to reflect a number of complex and interlinking factors such as confidence, classroom experience and motivation.

The most robust evidence on sex differences in mathematical competency can be derived from research in selected samples of mathematical gifted individuals. The widely cited 'Study of Mathematically Precocious Youth' (SMPY) conducted by Benbow and Stanley (1980) explored sex differences in mathematical reasoning ability, primarily assessed via performance on the American Scholastic Achievement Test-Mathematics (SAT-M). In a sample of 9,927 mathematically precocious 12-14 year olds, findings revealed a notable preponderance of males in the upper end of the distribution, with over 12 boys to every girl in the top 1% of scores. In a follow-up study of a further 40,000 intellectually precocious adolescents, these findings were confirmed (Benbow & Stanley, 1983). While a number of authors suggest that the preponderance of males stems from the variability of male scores on the SAT-M and other similar tests, the evidence suggests that mean scores on the SAT-M also favour boys by a consistent margin.

Further evidence for potential sex differences in numerical and mathematical competence may be derived from research regarding developmental dyscalculia. Developmental dyscalculia has been defined as a learning disability in mathematics the diagnosis of which is established when arithmetic performance in substantially below that expected for age, intelligence and education (American Psychiatric Association (APA), 1994). Despite some discrepancy with regard to definitional consistency the majority of evidence originating from a range of various countries identifies the prevalence of developmental dyscalculia to be approximately 3-6.5% in the normal school aged population (Badin 1983; Kosc 1974; Lewis et al 1994), a figure similar to that of dyslexia and attention deficit hyperactivity disorder (ADHD). Interestingly however, one of the most striking characteristics of developmental dyscalculia is that, unlike alternative learning disabilities such as dyslexia and autism in which males suffering from the disorder typically out-number those of females at least threefold (APA,1994), the prevalence rates of girls and boys suffering from developmental dyscalculia have been identified as relatively equal (Gross-Tsur et al., 1996). Equal numbers of males and females with dyscalculia therefore may actually indicate a male advantage in mathematics in comparison to other cognitive domains.

Overall, the research outlined above provides fairly convincing evidence for the presence of certain sex differences in mathematical ability. While these differences are undoubtedly task, sample, and age dependent, this pattern of results is similar to those revealed with regard to spatial cognition. Sex differences in spatial ability, like mathematical ability, rely heavily on the type of task being assessed. While a male advantage is typically reported in mental rotation and targeting (Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995) a female advantage has been reported for object location recognition and memory (Eals & Silverman, 1994; McBurney et al., 1997). Where sex differences favouring males do exist in spatial ability there is evidence that they may emerge on different tasks at different ages with their magnitude reported to increase with age (Collins & Kimura, 1997; Shute et al., 1983) and task difficulty (Prinzel & Freeman, 1995; Voyer et al., 1995). Interestingly however, despite their similarities, the potential impact of testosterone on sex differences in spatial ability has received considerably more attention than the potential impact of testosterone on similar sex differences in mathematical ability.

Related to evidence that precocious mathematical ability is more prevalent in males, it has also been hypothesised that PT may present one factor in the expression of exceptional mathematical talent via its facilitating influence on right hemispheric growth (Benbow & Stanley 1983; Kolata 1983). Based on the theoretical speculations of Geshwind and Galaburda (1987) that PT may slow the maturation of the left hemisphere resulting in concomitant compensatory growth in regions of the right hemisphere, evidence for the hypothesis is derived from research reporting a higher incidence of left handedness in mathematically gifted samples, and studies suggesting enhanced development of the right hemisphere and increased reliance on its capacities in the mathematically precocious.

Mathematically talented students are reported to be more than twice as likely as a comparative group of above average mathematical ability students to be left handed (Benbow, 1986, 1988). Higher frequencies of left-handers amongst university mathematics teachers and students have also been identified (see Benbow 1988). In keeping with the tenets of Geshwind and Galaburda's (1987) theory of cerebral dominance, as well as a higher incidence of left-handedness, Benbow (1988) also identified increased rates of myopia, incidence of allergy, migraine and other immune disorders in the mathematical precocious, all physiological correlates of PT exposure (as suggested by Geshwind & Galaburda, 1987).

Evidence for enhanced development and subsequent processing reliance on the capacities of the right hemisphere in the mathematically gifted comes from a series of studies conducted by O'Boyle and colleagues (e.g. O'Boyle & Benbow, 1990; O'Boyle O'Boyle, Alexander, & Benbow, 1991; O'Boyle, Gill, Benbow, & Alexander, 1994). In the study by O'Boyle & Benbow (1990) for example, a group of mathematically gifted and average ability adolescents performed a verbal dichotic listening task and a free-vision chimeric face task (involving the judgement of emotions). Results demonstrated the prototypic left hemisphere advantage to the processing of linguistic information on the verbal dichotic listening task in average ability adolescents, but no evidence for hemispheric asymmetry in mathematically gifted adolescents (they failed to show the usual left hemispheric advantage) on the same task. On the chimeric face task, while the typical right hemispheric bias for the processing of emotional information was revealed in both the average and mathematically gifted adolescents, the extent of right hemispheric bias in the mathematically gifted was appreciably stronger than that revealed for those with average ability.

Despite evidence for sex differences in certain aspects of mathematical competency and long-standing speculation as to a link between PT and subsequent mathematical abilities, possible associations and the nature of any potential interactions between the two are profoundly understudied. The paucity of previous research is surprising considering that, in Western society, the development and maintenance of numerical and mathematical competency is an educational priority. Evidence suggests that the possession of better mathematical ability and numeracy skills has a large positive impact on earnings and employability rates (Grogger & Eide, 1995; McIntosh & Vignoles, 2001). Further insight into the potential relationship between a biological influence, such as PT exposure, numerical and/or mathematical ability would increase our knowledge of the possible aetiological factors which may impact upon mathematical skill and possibly facilitate the delineation of biological and alternative factors which may contribute individual differences in mathematical ability. Given the reported value of mathematics, any information that may broaden our understanding of issues that may affect the development and expression of mathematical competency should be of paramount importance.

1.11 Evidence for an effect of prenatal testosterone on mathematical ability

To date, only a handful of studies have attempted to directly explore the possible relationships between PT exposure and mathematical abilities. A number of studies in individuals with CAH have noted deficits in quantitative skills on the Wechsler Adult Intelligence Scale (WAIS) Arithmetic subtest, the Arithmetic subtest of the Metropolitan Achievement Tests, and the Primary Mental Abilities Numerical subtest (Baker & Ehrhardt, 1974; Perlman, 1973; Sinforiani et al., 1994), actually suggesting that increasing exposure to testosterone may impair certain aspects of mathematical ability. Similarly, in children Finegan et al. (1992) reported a negative relationship between amniotic fluid testosterone levels and performance on counting and number fact tasks in girls aged 4; such that, high testosterone levels were associated with poorer performance on these tasks. However, no significant associations were revealed in boys.

Opposing evidence comes from Kimura and Carson (1995) who reported increased scores on a mathematical reasoning task (chosen as a male favouring assessment) in individuals with a R>L (male typical/higher testosterone) pattern of dermal ridge asymmetry. Similar findings are reported by Luxen and Buunk (2005) who revealed a significant negative correlation between right hand 2D:4D and numerical skills (assessed by scores on the Numerical Operations subscale of the Differential Aptitude Test) in both males and females, such that lower 2D:4D (higher PT) was associated with higher scores on the numerical assessment. Kempel et al. (2005) also found that females with low 2D:4D (masculinised), performed better on a numerical IQ task (continuing numerical series task) as compared with females showing a high 2D:4D (feminised) ratio, again suggesting a facilitative influence of PT exposure on numerical ability in females. Kempel et al. (2005) however reported no significant association in males.

A study by Fink, Brookes, Neave, Manning, & Geary (2006) revealed significant negative correlations between 2D:4D and number knowledge, counting, and visual number representation in boys aged 6-11, suggesting a facilitative influence of PT on such skills. Contrary to evidence from Luxen and Buunk (2005) and Kempel et al. (2005) however, no significant relationships were present in females. Similar findings were revealed by Brosnan (2008) who reported a significant negative association between 2D:4D and UK SAT numeracy scores in males aged 7, such that lower 2D:4D (higher testosterone) was associated with higher numerical SAT scores. No significant associations between 2D:4D and numeracy were revealed in females.

In summary, evidence to date generally implicates a potential association between levels of PT and certain aspects of numerical and mathematical cognition. Inconsistencies however exist with regards to both the significance of reported effects dependent upon sex, and the nature of possible relationships in females (where association have been reported in males the relationship is always positive, i.e. correlates of higher testosterone are associated with higher numerical/mathematical scores). Notably, sex differences in the presence of significant findings for an association between PT and cognition are also evident in the literature on spatial ability (Grimshaw, Sitarenios, & Finegan, 1995; Moffat & Hampson, 1996). Such findings therefore may generally reflect evidence for sex-dependent effects of testosterone on certain aspects of cognition.

A lack of consistency in previous findings is likely to relate to the fact that, even in very young children, numerical and mathematical skills are known to be based on many varied anatomically and functionally distinct cognitive components (Colvin, Funnell, & Gazzaniga, 2004; see Dehaene, 2000; and Geary, 2004) including spatial ability (e.g. Casey, Nuttall, Pezaris, & Benbow, 1995), verbal and linguistic abilities (e.g. Delgado & Prieto, 2004; Spelke & Tsivkin, 2001), and working memory (Bull, Johnston, & Roy, 1999; Bull & Sherif, 2001). The link between mathematical and spatial ability is often cited as one factor which may influence revealed sex differences in mathematics (Casey et al., 1995; Geary, 1996). Evidence from Casey et al. (1995)

reports the statistical removal of mental rotation ability to reduce sex differences on the SAT-M.

Geary (1996) argues that one of the primary factors which may account for sex differences in certain mathematical domains may be the presence of sex differences in certain spatial abilities, which may have emerged over the course of human evolution via the actions of intra-male sexual selection (the process by which traits evolve in response to same-sex competition). According to Geary's theory, the possession of superior navigational and tracking abilities in males may have facilitated superior performance during activities such as hunting, group migration and/or warfare leading to a reproductive advantage over their low-achieving peers, and ultimately an evolved sex difference in certain spatial abilities (in particular those involving the processing of 3-dimensional information). The link between spatial and mathematical ability however is not consistently reported. A meta-analysis by Friedman (1995), for example, presents only a slight correlation between the two, and actually suggests a greater correlation between mathematics and verbal ability. Geary (1996) himself recognises that the association between spatial ability and mathematical performance is complex and inconsistent. Ultimately, the extent to which sex differences in spatial ability may predict sex differences in mathematical ability is likely to be largely dependent on the nature of the specific task under observation in terms of both, the extent to which the mathematical task in question emphasises spatial representation in order to facilitate it's solution, and magnitude and significance of sex differences in the required aspect spatial ability depending on the task in question.

In addition to the many cognitive correlates of mathematics, it is also widely recognised that mathematical ability is influenced by a range of social and environmental factors including: parental expectation and instruction (Bleeker & Jacobs, 2004; Muller, 1998), education (Reusser, 2000) and stereotype threat, (e.g. Ben-Zeev, Fein, & Inzlicht, 2005; Cadinu, Maass, Rosabianca, & Kiesner, 2005). Such factors may also go on to influence or exaggerate sex differences in mathematical ability, for example, research exploring the impact of stereotype threat on women's mathematical performance shows that when a mathematical task is explicitly characterised as sensitive to sex differences women significantly underperform in relation to men (Spencer et al., 1999). While a full discussion and dissection of the many variables associated with mathematical ability is beyond the scope of this thesis it is vital to recognise that mathematics is not a univariate construct, and that marked individual differences have been recognised in most of the multiple components (social and cognitive) thought to contribute to mathematical competence and performance.

1.12 Aims and Hypothesis of the thesis

Despite speculation regarding a potential association between PT and mathematical skill relatively few studies have attempted to explore any correlation between the two. Where possible relationships have been investigated no two studies have employed the same numerical or mathematical measures. Given the multifaceted nature of mathematical competency this makes the outcomes of the different experiments conducted to date difficult to compare. The widely different numerical and mathematical assessments adopted in previous research across studies may offer one possible explanation for the discrepancy in findings across different experiments. To date no one experiment or series of studies have attempted to provide a thorough and methodical investigation of the possible relationship between PT and numerical or mathematical skill.

The purpose of this thesis is therefore to systematically explore any potential relationship between 2D:4D, as a proxy for PT exposure, and mathematical and numerical skills. Given the reported importance of numerical and mathematical competence in Western society any factor which may impact upon such skills requires serious consideration. The findings of the current thesis will also contribute to an expanding body of evidence regarding the potential cognitive correlates of PT effects (and 2D:4D).

Chapter 2

Experiment 1: 2D:4D and Numerical Competence in Children.

2.1 Introduction

There is longstanding speculation as to a potential association between exposure to prenatal testosterone (PT) and subsequent mathematical abilities, research directly pertaining to this relationship however is relatively scarce. Evidence from the limited range of research that has been conducted, (Baker & Erhardt, 1974; Brosnan, 2008; Finegan et al., 1992; Fink et al., 2006; Kempel et al., 2005; Kimura & Carlson, 1995; Luxen & Buunk, 2005; Perlman, 1973; Sinforiani et al., 1994) presents equivocal support with regard to the nature of reported effects depending upon sex (several authors report significant effects in one sex only) and the direction of revealed relationships in females (where significant effects have been reported in males the association is always positive, i.e. higher numerical/ mathematical scores are related to correlates of higher PT), see chapter 1. Research to date however has employed a variety of widely different mathematical and numerical assessments in order to investigate potential correlations making comparisons across studies difficult. Furthermore, the majority of research conducted thus far was not designed specifically to test correlations between PT and mathematical ability and/or has investigated relations on only one or two numerical or mathematical tasks, thus providing limited information and consideration of any possible association between PT and numerical or mathematical skill.

Evidence suggests that children's knowledge of certain aspect of quantity and basic arithmetic may be an inherent and potentially domain-specific cognitive ability (Dehaene, 1997; Gallistel & Gelman, 1992; Geary, 1995). A wealth of empirical literature now exists which demonstrates evidence for elementary number discrimination and computation in infants (e.g. Antell & Keating, 1983; Sharon & Wynn, 1998; Starkey, Spelke, & Gelman, 1990; Wynn, 1992; Xu and Spelke, 2000) and a variety of animal taxa (e.g. Brandon & Terrace, 1998; Church & Meck, 1984; Hauser, Tsao, Garcia, & Spellke, 2003). In addition, a convincing body of research based upon brain imaging studies and the assessment of patients with various neurological and developmental deficits, presents evidence that numerical processing may be sub-served by a distinct neural circuitry (Dehaene, Piazza, Pinel, & Cohen, 2003). Based on the rationale that any biological influence may be stronger on basic numerical abilities that potentially tap inherent numerical skills Fink et al. (2006) employed the Number Processing and Calculation in Children (NUCALC) Test Battery (Deloche et al., 1995; von Aster, 2001) in order to explore correlations between basic numerical proficiency and the second to fourth digit ratio (2D:4D) as a potential proxy of PT exposure in a sample of British and Austrian children aged 6-11 years. The NUCALC consists of 11 subtests from which four dimensions of mathematical competency can be assessed; general mathematical ability (represented the child's overall score on the test battery), number knowledge, counting and visual representation. After controlling for age and ethnicity Fink et al. (2006) reported significant negative correlations between right and left hand 2D:4D and; total scores on the NUCALC test battery, number knowledge scores, and counting scores in males. A significant negative association between 2D:4D (low 2D:4D = high PT) and visual representation scores in males was also found but for right hand ratios only, although the relationship between left hand 2D:4D and visual representation in males was approaching significance (also in a negative direction). No significant association between 2D:4D and scores on the NUCALC were revealed in females. To the extent that 2D:4D reflects levels of testosterone exposure *in utero* the findings suggested that high PT was associated with improved mathematical performance in males. Other research in children aged 7 (Brosnan, 2008) reported similar findings to Fink et al. (2006) for scores on UK Standardized Assessment Tests (SAT), with significant negative correlations identified between 2D:4D and numerical performance in males and no significant correlation between 2D:4D and numerical performance in females. Further replication of the findings of Fink et al. (2006) would therefore offer strong support for a potential role of PT on basic numerical and mathematical competencies in males.

The NUCALC test battery utilised by Fink et al. (2006) is primarily a diagnostic instrument, aimed at a fairly wide developmental age range (6-11 years old) for the identification of children with possible mathematical difficulties (dyscalculia). Ultimately therefore the battery offered relatively limited scope for the evaluation of various numerical abilities which may be associated with 2D:4D. The present study aimed to replicate and extend research by Fink et al. (2006) by exploring associations between 2D:4D and performance on a similar but more generalised battery of basic numerical tasks designed specifically for the current study. Using national curriculum guidelines, current educational and psychological literature and various standardised assessments of mathematical difficulties, two comparable though separate test batteries were designed to assess basic numerical and mathematical skill in two distinct age ranges. One test battery was developed for children aged 5-7, and a second battery was developed for children aged 8-11. Two separate test batteries were created so that associations could be evaluated across the primary school age range without compromising the appropriateness of assessment according to age. Based on previous evidence for a potential association between potential indices of PT exposure and aspects of numerical and mathematical performance it was hypothesised that 2D:4D would be associated with performance on the mathematical test battery. In light of evidence from both Fink et al. (2006) and Brosnan (2008) conducted with children of a similar age range, it was expected that lower 2D:4D (higher PT) would be associated with higher scores on the test battery in males. Given the contradictory results of previous research in females no clear predictions could be made with regard to potential findings in girls. Based on prior evidence however it was hypothesised that any revealed associations between 2D:4D and performance on the mathematic test battery may differ (in terms of the direction and/or strength) depending upon sex. In light of previous evidence which generally suggests a lack of sex differences in basic mathematical competency (Fink et al., 2006; Geary, 1996), no significant sex differences mathematical performance were anticipated.

2.2 Method

2.2.1 Design

The study employed a correlational design in order to explore the relationship between 2D:4D and performance on the mathematical test battery in both males and females.

2.2.2 Participants

Seventy-nine participants (41 males; 38 females) were recruited from four North East primary and first schools. Participants were recruited on a voluntary basis, subject to full informed, written school and parental consent. Based on evidence that handedness may be related to PT and, more specifically, the somatic marker 2D:4D (see chapter 1) data from left handed participants (assessed according to writing hand) was omitted from the analysis (4 males; 2 females). In total, data recruited from 19 males (mean age = 5.84; SD = 0.37) and 22 females (mean age = 5.86; SD = 0.36) aged 5-7 years, and 18 males (mean age = 9.29; SD = 0.56) and 14 females (mean age = 9.65; SD = 0.52) aged 8-11 years were included in the analysis. The parents of participating children provided information regarding their child's date of birth, any potential past or previous injury to the second or fourth finger, and any known hormonal abnormalities. According to this information no participants possessed any hormonal abnormalities or any injury to the second or fourth fingers of either hand.

2.2.3 Measures

2.2.3.1 Numerical test battery

Two version of a numerical test battery were developed, one aimed children at aged 5-7 years old and a second aimed at children aged 8-11 years old. Both versions consisted of sixteen mathematical tasks which could be further collapsed into seven broad categories of basic mathematical skill (see table 1). The battery designed for children aged 8-11 years old was simply an extension of the battery designed for children aged 5-7 years old such that for the majority of tasks children aged 8-11 years completed a greater number of and slightly more demanding examples of the same tasks. All task instructions were read aloud to the participating children, each task was also preceded by an example described to the children by the experimenter. Children were required to provide either an oral or written response (customized workbooks were provided for written responses). A brief description of each sub task is given below. Copies of the instructor booklets (including details regarding scoring for each subtask) can be found in appendices 8 and 9, appendix 10 contains example stimuli for any tasks involving visual presentation of information.

Task 1 – Counting: Children were presented with various arrays of dots or pictures and requested to count items aloud while simultaneously pointing to each dot/picture. Children aged 5-7 years were presented with five dot/picture arrays while children aged 8-11 years were presented with ten dot/picture arrays.

Task 2 - Reading numbers: Children were asked to recite aloud visually presented Arabic numerals. Ten numerals in the range of 2-25 were presented to children aged 5-7 years. Children aged 8-11 years were presented with twenty written numerals ranging from 2-63002.

Task 3 - Writing numbers: Children were requested to write in Arabic numerals a series of orally dictated numbers. Children aged 5-7 years were presented with 5 numbers ranging from 3-14 while children aged 8-11 years were presented with ten numbers in the range of 3-4685

Tasks 4,5,6 - Mental arithmetic: Mental calculation tasks were orally presented and consisted of standard addition (task 4), subtraction (task 5) and everyday addition and subtraction e.g. "*There are three people on the bus. One more person gets on how many are now on the bus?*" (task 6) problems. Children aged 5-7 years were presented with 3 mental additions, 3 mental subtractions and 4 everyday addition and subtraction numerical problems while children aged 8-11 years were presented with 5 mental additions, 5 mental subtractions and 5 everyday numerical addition and subtraction problems.

T*asks 7,8,9,10,11,12 – Number line tasks:* Firstly children were required to identify the exact numerical value towards which an arrow was directed on a number line (task 7). Children aged 5-7 years were presented with two number lines ranging from 0-10 while children aged 8-11 years were presented with two number lines ranging from 0-10, one number line ranging from 0-100 and one from 0-1000. Each number line was clearly and equally divided into ten sections. Arrows were only directed at dividing lines. Children were then presented with blank number lines (task 8), i.e. lacking any dividing lines, and asked to estimate the number towards which an arrow was pointing. Again children aged 5-7 years were presented with two number lines raging from 0-10, while children aged 8-11 years were presented with four number lines all similarly ranging from 0-10. Arrows were only ever directed at whole numbers. Following this (task 9), further blank number lines were presented (two to children aged 5-7 years ranging from 0-10, four to children aged 8-11 years two ranging from 0-10 and two from 0-100) and children were requested to make a small mark on the line where they believed a particular number to be positioned. Numbers were orally dictated. In task 10 children were sequentially presented with various number sequences (younger children with a total of two sequences, older children with a total of four) containing missing elements and asked to identify the position of an absent numeral. All numbers were orally dictated. Children were then shown various sets of numbers presented in a random order (task 11) and asked to rewrite the number in the correct numerical order beginning with the smallest. Children aged 5-7 years were presented with two random number sets and children aged 8-11 years with four. Finally, (task 12) children were presented with various sets of ten pictures (again children aged 5-7 years with two sets and children aged 8-11 years with four) positioned in a straight line. Children were orally dictated a serial position, e.g. fourth, and required to point to the picture in that particular position.

Tasks 13,14,15 – Number comparison: In task 13 children were sequentially presented with two large circles containing an unequal number of dots and requested to identify (via pointing) the circle with the greatest number of dots as quickly as possible. Following this (task 14) children were again presented with two large circles, one containing a written Arabic numeral and the other a series of dots and asked to state whether or not the written Arabic numeral corresponded to the number of dots in the adjacent circle. Finally, in a written number comparison task (task 15) pairs of numbers were presented to children as Arabic numerals and children were required to circle the greatest number from each pair. For each number comparison subtask children aged 5-7 years were presented with a total of five comparisons and children aged 8-11 years with a total of ten comparisons.

T*ask 16 – Estimation:* Both younger and older children were presented with 5 arrays of various pictures and asked to orally approximate the number of items.

Table 1

The mathematical sub categories and component individual tasks (and associated scoring) included in both versions of the test battery.

Category	Task (in order of presentation)	Max. score	
		5-7 yrs	8-11yrs
Counting	Counting	10	20
Reading numbers	Reading numbers	10	20
Writing numbers	Writing numbers	10	20
Mental arithmetic	Addition	6	10
	Subtraction	6	10
	Everyday addition and subtraction	8	10
Number line tasks	Identification of a number on an analogue scale	2	4
	Approximate identification of number on an analogue scale	4	8
	Approximate positioning of number on analogue scale	4	8
	Identification of the missing number in a numerical series	2	4
	Arranging numbers	2	4
	Positioning numbers	2	4
Number comparison	Quantity comparison	5	10
	Arabic digit-quantity comparison	5	10
	Arabic digit comparison	4	8
Estimation	Estimation	10	10

2.2.3.2 Second to fourth finger ratio measures

Direct measurements of the second and fourth fingers were taken from the basal crease to the proximal tip on the ventral surface of both left and right hands using Vernier Callipers accurate to 0.01mm (see figure 1). In order to ensure repeatabilities measurements were taken twice, once prior to conducting the experiment and once following completion. An average length of 2^{nd} and 4^{th} fingers was calculated and 2D:4D computed using these averaged measurements. Intraclass correlation coefficients (r_I) showed high retest-reliability between first and second measurements of both the right (second $r_I = 0.975$; fourth $r_I = 0.971$; 2D:4D $r_I = 0.735$) and left hands (second $r_I = 0.976$; fourth $r_I = 0.98$; 2D:4D $r_I = 0.78$). From initial and final second and fourth finger measurements the technical error of measurement (TEM) and relative technical error of measurement (rTEM) were computed according to protocol established by Weinberg, Scott, Neiswanger and Marazita (2005). For second digit measurements TEM was computed to be 0.9 and 0.86 with the rTEM calculated to be 1.68% and 1.59% for the right and left hands respectively. TEM for fourth digit measurements was calculated at 0.95 for right hand measures and 0.8 for left hand measures with corresponding rTEM calculated to be 1.74% and 1.45% respectively. According to Weinberg et al. (2005) smaller TEM and rTEM values represent more precise measurements. With regard to rTEM scores, Weinberg et al. (2005) recommends a cut-off point of 5% with all rTEM percentages above this considered imprecise. According to this published criterion, an acceptable degree of precision for second and fourth finger measurements was met.

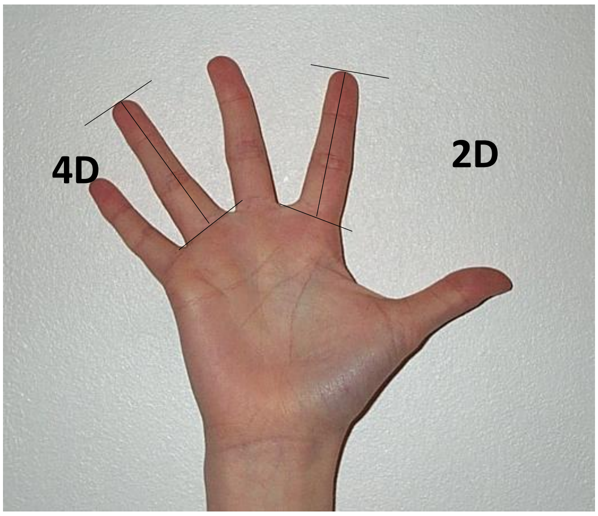


Figure 1. Second and fourth digit measurements, from the basal crease to the proximal tip on the ventral surface of the hand. 2D:4D = length of the second finger (2D) ÷ length of the fourth finger (4D).

Contrary to expectation no significant sex differences in 2D:4D values were identified (left hand $t_{(71)} = 0.005$, $p = 0.996$; right hand $t_{(71)} = 0.267$, $p = 0.79$). Furthermore, while mean left hand 2D:4D values were equal across males and females (male mean = 0.983, SD = 0.037; female mean = 0.983, SD = 0.04), in direct opposition to previous evidence, mean right hand 2D:4D values were actually revealed to be marginally lower in females (mean = 0.98, SD = 0.04) as compared to males (mean = 0.983, SD = 0.037).

2.2.4 Procedure

The study was approved by Northumbria University, School of Psychology and Sport Sciences Ethics Committee. Following full written school and parental consent as well as oral assent from the child, the participants were individually assessed in a quiet room. Testing took approximately 20-30 minutes per child depending upon age and ability. Children were first given a full verbal brief and an initial finger measurement on the ventral surface of both the right and left hands were collected. According to the age of the child the appropriate test battery was then completed. Following task completion a second 2D:4D measure was taken. Children were finally given a full verbal de-brief. Participation was completely voluntary, no form of participant payment was offered.

2.3 Results

In order to standardise the data in the current study, raw scores on the numerical test batteries were converted into percentages. As the administered test battery and subsequent scoring differed depending on whether the child was aged 5-7 or 8-11 years old data from these two sets of children were analysed separately. One-sample Kolmogorov-Smirnov tests were used in order to explore normality. The results of these analyses revealed that scores on the numerical test battery were not normally distributed (see Appendix 1) thus two-tailed Spearman correlation coefficients (ρ) were used to assess the relationship between 2D:4D and numerical performance. Mann-Whitney U tests were adopted in order to analyse sex differences in performance. Full tables of means for right and left hand 2D:4D values and performance on the numerical test battery in both males and females separately and the sample overall can be viewed in Appendix 1.

2.3.1 2D:4D and performance on the numerical test battery

Table 2 displays the ρ correlations between 2D:4D and performance on the numerical test battery for the total sample and males and females analysed separately in children aged 5-7 years. No significant correlations were found between either right or left hand 2D:4D and performance in the overall sample (male and females data combined). A significant positive correlation however was identified between right hand 2D:4D and number comparison scores in female data analysed separately, thus higher 2D:4D (lower PT) was associated with higher scores in the number comparison sub-category. No further significant associations were observed in females and no significant relationships between either right or left hand 2D:4D and performance were identified in male data analysed independently.

Table 3 shows ρ correlations between 2D:4D and performance on the numerical test battery in children aged 8-11 years. No significant relationships between either right or left hand 2D:4D and performance on the numerical test battery were observed in the overall sample or males and females analysed independently in this age group.

Retrospective power analysis was conducted for all computed correlations. According to Siegal and Castellan (1988) the power of the Spearman rank-order correlation when compared to the parametric equivalent Pearson product-moment correlation is approximately 91%, thus the Spearman's correlation in a population of 100 cases would demonstrate similar significance as a Pearson's correlation in a population of 91 cases. Based on this calculation adopted sample sizes were adjusted to take into account the relative loss of power for the Spearman's analyses relative to a Pearson's (sample size multiplied by 0.91), power for each analyses was then calculated using G*Power 3.1.2 (Faul, Erdfelder, Buchner, & Lang, 2009). A similar procedure for adjusting sample sizes is proposed by Clark-Carter (2010) who suggests that when using a Spearman rank-order correlation, in order to achieve the equivalent level of power to that which would be observed with a Pearson's the sample size should be multiplied by 1.1 (the inverse of 0.91). As can been seen in tables 2 and 3, power was generally very low.

Given that power was low it is important to highlight that although only one significant correlation was revealed, a number of the correlations coefficients from the analysis suggest small to moderate and even moderate to high effect sizes for analysis of both younger and older children. Interestingly however the direction of the relationships for such correlation were not consistent across tasks within each group (males, females and overall). In females aged 8-11 years for example, while a moderate to large effects size was observed between 2D:4D and number comparison task performance in a positive direction (although non-significant), moderate to large effect sizes were also observed for counting task performance (again non-significant) but in a negative direction.

Table 2

Spearman correlation coefficients (ρ) for the relationship between left and right hand 2D:4D and performance on the numerical test battery in children aged 5-7 years old. P values and power calculations (1- β) are also listed. Significant results are highlighted in bold.

	Males and Females n = 41		**Males n = 19**		**Females n = 22**	
	Right 2D:4D	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Counting	ρ = -0.103 p = 0.521 1-β = 0.094	ρ = -0.196 p = 0.219 1-β = 0.219	ρ = -0.026 p = 0.915 1-β = 0.051	ρ = -0.16 p = 0.512 1-β = 0.097	ρ = -0.143 p = 0.526 1-β = 0.092	ρ = -0.268 p = 0.228 1-β = 0.209
Reading numbers	ρ = -0.154 p = 0.337 1-β = 0.152	ρ = -0.203 p = 0.202 1-β = 0.233	ρ = -0.201 p = 0.41 1-β = 0.125	ρ = -0.133 p = 0.587 1-β = 0.082	ρ = -0.124 p = 0.583 1-β = 0.081	ρ = -0.242 p = 0.278 1-β = 0.178
Writing numbers	ρ = -0.097 p = 0.548 1-β = 0.089	ρ = -0.166 p = 0.299 1-β = 0.169	ρ = -0.091 p = 0.711 1-β = 0.065	ρ = -0.097 p = 0.694 1-β = 0.067	ρ = -0.071 p = 0.755 1-β = 0.06	ρ = -0.23 p = 0.302 1-β = 0.164
Mental arithmetic	ρ = 0.076 p = 0.637 1-β = 0.074	ρ = 0.052 p = 0.745 1-β = 0.061	ρ = 0.08 p = 0.746 1-β = 0.061	ρ = 0.131 p = 0.592 1-β = 0.081	ρ = 0.092 p = 0.684 1-β = 0.067	ρ = 0.019 p = 0.934 1-β = 0.051
Number line tasks	ρ = -0.116 p = 0.471 1-β = 0.106	ρ = 0.163 p = 0.308 1-β = 0.165	ρ = -0.041 p = 0.868 1-β = 0.053	ρ = 0.37 p = 0.118 1-β = 0.336	ρ = -0.203 p = 0.364 1-β = 0.137	ρ = 0.025 p = 0.912 1-β = 0.051
Number comparison	ρ = 0.261 p = 0.1 1-β = 0.358	ρ = 0.111 p = 0.49 1-β = 0.101	ρ = -0.072 p =0.77 1-β = 0.059	ρ = -0.126 p = 0.607 1-β = 0.078	**ρ = 0.568** **p = 0.006** **1-β = 0.779**	ρ = 0.321 p = 0.145 1-β = 0.286
Estimation	ρ = -0.081 p = 0.615 1-β = 0.077	ρ = -0.02 p = 0.903 1-β = 0.052	ρ = 0.004 p = 0.988 1-β = 0.05	ρ = 0.392 p = 0.097 1-β = 0.375	ρ = -0.14 p = 0.535 1-β = 0.09	ρ = -0.293 p = 0.186 1-β = 0.244
Overall	ρ = -0.014 p = 0.931 1-β = 0.051	ρ = 0.005 p = 0.974 1-β = 0.05	ρ = -0.056 p = 0.819 1-β = 0.055	ρ = 0.231 p = 0.341 1-β = 0.152	ρ = 0.022 p = 0.922 1-β = 0.051	ρ = -0.174 p = 0.438 1-β = 0.113

Table 3

Spearman correlation coefficients (ρ) for the relationship between left and right hand 2D:4D and performance on the numerical test battery in males and females in children aged 8-11 years old. . P values and power calculations (1- β) for each analysis are also listed.

	Males and Females n = 32		**Males n = 18**		**Females n = 14**	
	Right 2D:4D	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Counting	ρ = 0.102 p = 0.578 1-β = 0.083	ρ = 0.169 p = 0.354 1-β = 0.145	ρ = 0.403 p = 0.097 1-β = 0.373	ρ = 0.422 p = 0.081 1-β = 0.408	ρ = -0.447 p = 0.109 1-β = 0.347	ρ = -0.378 p = 0.182 1-β = 0.252
Reading numbers	ρ = -0.016 p = 0.932 1-β = 0.051	ρ = -0.159 p = 0.384 1-β = 0.134	ρ = -0.003 p = 0.99 1-β = 0.05	ρ = -0.181 p = 0.472 1-β = 0.106	ρ = 0.007 p = 0.981 1-β = 0.05	ρ = -0.11 p = 0.708 1-β = 0.064
Writing numbers	ρ = -0.078 p = 0.673 1-β = 0.069	ρ = -0.176 p = 0.337 1-β = 0.154	ρ = 0.19 p = 0.451 1-β = 0.112	ρ = -0.064 p = 0.802 1-β = 0.057	ρ = -0.284 p = 0.325 1-β = 0.156	ρ = -0.221 p = 0.448 1-β = 0.111
Mental arithmetic	ρ = 0.133 p = 0.466 1-β = 0.108	ρ = 0.112 p = 0.542 1-β = 0.09	ρ = -0.061 p = 0.81 1-β = 0.056	ρ = -0.089 p = 0.725 1-β = 0.063	ρ = 0.138 p = 0.639 1-β = 0.073	ρ = 0.309 p = 0.283 1-β = 0.177
Number line tasks	ρ = 0.26 p = 0.151 1-β = 0.287	ρ = 0.171 p = 0.349 1-β = 0.147	ρ = 0.367 p = 0.134 1-β = 0.313	ρ = 0.313 p = 0.207 1-β = 0.234	ρ = 0.133 p = 0.651 1-β = 0.071	ρ = 0.046 p = 0.875 1-β = 0.052
Number comparison	ρ = -0.068 p = 0.712 1-β = 0.065	ρ = -0.043 p = 0.814 1-β = 0.056	ρ = -0.395 p = 0.105 1-β = 0.359	ρ = -0.286 p = 0.25 1-β = 0.201	ρ = 0.426 p = 0.129 1-β = 0.316	ρ = 0.486 p = 0.078 1-β = 0.411
Estimation	ρ = -0.089 p = 0.629 1-β = 0.075	ρ = -0.065 p = 0.724 1-β = 0.063	ρ = -0.006 p = 0.98 1-β = 0.05	ρ = -0.098 p = 0.7 1-β = 0.066	ρ = -0.181 p = 0.536 1-β = 0.09	ρ = -0.077 p = 0.794 1-β = 0.057
Overall	ρ = 0.049 p = 0.791 1-β = 0.057	ρ = -0.075 p = 0.685 1-β = 0.068	ρ = -0.128 p = 0.612 1-β = 0.077	ρ = -0.234 p = 0.35 1-β = 0.147	ρ = 0.095 p = 0.748 1-β = 0.061	ρ = 0.099 p = 0.736 1-β = 0.062

2.3.2 Correlations with age

Within each age group (5-7 and 8-11 years) Spearman's correlations were used to assess any potential relationship between age and; overall performance on the numerical test battery and performance on each subcategory of the numerical test battery. No significant correlations between age and performance were revealed for children aged 5-7. Significant correlations were revealed however between age and; writing numbers scores, r = 0.357, p = 0.045, and scores on the number line tasks subcategory, r = 0.419, p = 0.017 in children aged 8-11 years (a full list of ρ and p values relating to analysis of the relationship between age and performance can be viewed in Appendix 1). Both significant correlations were in a positive direction, demonstrating improved performance with increasing age.

2.3.3 Sex differences

Mann-Whitney *U* tests revealed no significant sex differences in performance on the numerical test battery in children ages 5-7 years old. Significant sex differences were revealed in children aged 8-11 years old for performance on the counting subtask, *U* = 85, Z =-2.059 , p = 0.039,with females scoring significantly higher on the task (mean = 99.29, SD = 2.67) in comparison to males (mean = 95.28, SD = 6.52). No further significant sex differences in performance however were identified. A full list of *U*, Z, and p values relating to the analyses for sex differences can be viewed in Appendix 1.

2.3.4 Reliability analysis

Simple visual inspection of the data suggests the presence of ceiling effects on the devised assessments. As can be seen in table 4 the large majority of average scores for the different sub categories of the test batteries were above 70%. A more detailed breakdown of the descriptive statistics relating to performance on each subtask of both test batteries can be found in Appendix 11.

Reliability analysis exploring the collapsibility of the tasks included in the mental arithmetic, number lines tasks and number comparison sub categories was also conducted. Internal consistency of the mental arithmetic subcategory as expressed by

Cronbach's alpha coefficient was satisfactory for the battery developed for children aged 5-7 years (0.789). Reliability analysis of the mental arithmetic subcategory for the battery developed for children aged 5-7 years however did not reach an acceptable degree of internal consistency (Cronbach's alpha = 0.686). Poor internal consistency was also revealed for analysis of both the number line tasks subcategory (Cronbach's alpha 5-7 year olds = 0.612; Cronbach's alpha 8-11 year olds = 0.459) and the number comparison subcategory (Cronbach's alpha 5-7 year olds = 0.524; Cronbach's alpha 8-11 year olds = 0.274).

Table 4

Average scores, standard deviations and minimum and maximum values for performance (overall and on each sub category) on the numerical test batteries in children aged 5-7 and 8-11 years.

Task	5-7 years old (n = 41)				8-11 years old (n = 32)		
	Mean	Min-Max	SD		Mean	Min-Max	SD
Counting	96.83	50-100	9.34		97.03	80-100	5.51
Reading numbers	91.46	10-100	19.69		91.41	55-100	10.57
Writing numbers	89.76	0-100	19.56		95.63	70-100	9.22
Mental Arithmetic	68.54	15-100	24.53		68.75	13.33-93.33	18.47
Number line tasks	68.45	25-93.75	14.11		79.46	3.13-100	17.17
Number comparison	81.88	50-100	13.66		91.96	75-100	7.09
Estimation	23.66	0-100	21.3		39.69	10-100	22.36
Test battery overall	73.69	46.67-90	11.13		83.32	52.5-93.13	8.42

2.4 Discussion

The aim of the current study was to replicate and extend research conducted by Fink et al. (2006) exploring any potential association between basic numerical abilities and 2D:4D as putative marker of PT. Utilising two versions of a numerical test battery designed specifically for the current study, one aimed at children aged 5-7 and the second aimed at children aged 8-11, the present experiment attempted to investigate associations utilising a wider range of numerical skills to those examined by Fink et al. (2006). The results revealed a significant positive correlation between right hand 2D:4D and number comparison sub-category scores in females aged 5-7 years old. As high 2D:4D is thought to indicate lower PT exposure the findings suggest a detrimental effect of PT exposure on number comparison performance in females aged 5-7 years. No significant correlations however were identified in the overall sample of children aged 5-7 years old or in males aged 5-7 years old analysed separately. Similarly no significant correlations between either right or left hand 2D:4D and numerical task performance were observed in children aged 8-11years old.

The finding of a significant correlation between right hand 2D:4D and counting scores in females in the absence of a similar effect in males suggests that the nature of PT effects may differ depending upon sex. This is in line with previous evidence where a number of authors have reported significant relationships between potential measures of PT and mathematical performance in one sex only (males only – Brosnan, 2008; Fink et al., 2006; females only – Finegan et al., 1992; Kempel et al., 2005). The effect however was not replicated in children aged 8-11 years. As you would expect any biological, organisational influence on cognition to be consistent across different ages ranges the discrepancy between findings for children aged 5-7 years and 8-11 years suggests that other factors may influence potential associations. As described in chapter 1, it is widely recognised that mathematical ability is influenced by a range of social and environmental factors such as parental expectation and instruction, education and stereotype threat. As you would expect ongoing exposure to such factors with increasing age it is possible that the influence of social and experiential factors may account for the discrepancy in findings between younger and older children. It is important to recognise however that both minimum and average scores on the counting tasks were notably high, implying possible ceiling effects on performance. In light of potential ceiling effects, speculation regarding the outcome of the findings should be considered with caution.

A similar significant correlation between 2D:4D and number comparison scores in females aged 5-7 years was not revealed when considering left hand 2D:4D data. This however is in line with previous evidence where associations between 2D:4D and behavioural and cognitive outcomes are commonly reported to be stronger, or only present, in the right hand (e.g. Brown, Finn, & Breedlove, 2002; Csathó et al., 2003; Lutchmaya et al., 2004; Williams et al., 2000). The precise reason for this remains unknown. More unusual is the failure to find significant sex differences in either right or left hand 2D:4D measures. There is now well established evidence for sex differences in 2D:4D with a wide range of previous research reporting significantly lower 2D:4D values in males as compared to females (e.g. George, 1930; Phelps, 1952; Manning et al., 1998; Manning, Barley, et al., 2000; Manning, 2002, see chapter 1). While failure to replicate significant sex differences is not unique to the present study (e.g. Brosnan, 2008) particularly unusual is the fact that identified mean 2D:4D values also failed to reflect the anticipated pattern of results. In the present experiment mean 2D:4D values in males and females were revealed to be equal for right hand measures and marginally lower in females for left hand 2D:4D measures. Crucially however, the current study did possess one important confound in that the ethnicity of the participating children was not recorded, and subsequently not controlled for. Large population and racial differences have been observed in both adults (Manning, Barley, et al., 2000; Manning, Henzi, Venkatramana, Martin, & Singh, 2003; Peters, Tan, Kang, Taixeira, & Mandal, 2002) and children (Manning, Stewart, Bundred, Trivers, 2004; McIntyre, Ellison, Lieberman, Demerath, & Towne, 2005) to the extent that ethnicity actually accounts for a greater proportion of the variation in 2D:4D than sex (McIntyre, 2006). Failure to control for ethnicity may thus result in spurious results caused by within sample ethnic variation. Importantly, the magnitude of sex differences in 2D:4D appears to be similar across different populations (Manning, Stewart, et al., 2004; Manning, Henzi, et al., 2003; Peters et al., 2007) suggesting that sex differences in the trait are independent of any ethnic variation. The problem of racial confounding therefore can be easily avoided with careful study design. All future studies conducted as part of the current thesis therefore will take into consideration any potential 2D:4D differences as consequence of ethnicity.

Research by Jordan, Kaplan, Olah, and Locuniak (2006) presented evidence for a significant male advantage in counting, number knowledge, non verbal calculation, estimation and pattern recognition. Previous findings by Jordan, Hanich and Kaplan (2003) also found a small advantage on estimation tasks while evidence from Carr & Jessup (1997) found evidence for a tendency for girls to use more language-based counting strategies in order to solve arithmetic problems and boys to have a small advantage on estimation tasks. Aside from such studies however the majority of previous research generally suggests that sex differences do not exist for basic mathematical skills (e.g. Geary, 1996; Fink et al., 2006). Contrary to both previous research and the pattern of results reported by Jordan et al. (2006) the current study revealed a significant female advantage for performance on the counting task in children aged 8-11 years old. If such sex differences have a purely biological origin you would expect similar sex differences to also emerge in children aged 5-7 years old. It is likely therefore that social and environment factors have, at the very least, contributed to the revealed sex differences in performance. Ultimately however given the potential ceiling effects on counting performance, speculation regarding sex differences on the task remains tentative.

Ceiling effects on the numerical assessment were not restricted to performance on the counting task. High average scores in conjunction with a limited range of scores focused at the upper end of the scale were similarly revealed for a number of other sub categories on the numerical test battery. As a consequence of such effects individual differences in numerical competency may not have been adequately evaluated. In addition, poor internal consistency was revealed for each subcategory that contained more than one tasks (i.e. mental arithmetic, number line tasks and number comparison). Poor internal consistency suggests that each of tasks employed within each of the three subcategories were actually measuring different constructs. Ultimately therefore little meaning can be derived from overall averaged scores on the mental arithmetic, number line tasks or number comparison sub categories. Such findings may be related to the revealed ceiling effects on performance. The lack of internal consistency on what appear to be very similar mathematical tasks may also reflect the exceptionally multifaceted nature of mathematical ability.

Besides the revealed issues regarding the adopted measure of numerical skill one further important limitation of the current study which should be recognised is sample size. Despite an initial sample size of seventy-nine, participant numbers were dramatically reduced upon segregation by age and sex. Unsurprisingly therefore calculated power for each analysis was typically low. Larger sample sizes are required in order to reach any firm conclusions regarding the nature of reported effects. Interestingly, although only one significant correlation was revealed a number of the correlation coefficients suggested the presence of effect sizes similar to those reported by Fink et al. (2006) and Brosnan et al. (2008), i.e. small to medium and some medium to large according to the conventions of Cohen (1988). Crucially however the nature of such effects were not consistent across the different sub-tasks included in the test battery within each group i.e. overall and males and females separately. While Fink et al. (2006) reported all negative correlation with 2D:4D in males across all numerical measures derived from the NUCALC test battery (see chapter 1), in the present experiment both positive and negative correlations were identified in both males and females across the different sub-tasks included in the test battery.

Future research should usefully address the methodological limitations of the current study. In order to address ceiling effects during the assessment of basic numerical skill the adoption of reaction time measures may be useful. While paper and pencil assessments of accuracy alone (such as that adopted in the current study and research by Fink et al., 2006) may be suitable for the diagnosis of selective mathematical deficits, given evidence that speed of processing may be an important predictor of mathematical performance (Bull & Johnston, 1997) the inclusion of reaction time measures is likely to provide a more sensitive tool for the identification of individual differences in normal populations, particularly on basic numerical tasks where accuracy is generally high.

As mentioned previously a growing body of evidence suggests that knowledge of certain aspect of quantity and basic arithmetic may be an innate. In an attempt to replicate Fink et al. (2006) the current study explored relations between basic numerical skill and 2D:4D in children, based on the rationale that any biological influence may be stronger on basic numerical abilities that potentially tap inherent numerical skills. The tasks employed in both the current study and the study conducted by Fink et al. (2006) however probably reflect a combination of both inherent knowledge and school-based learning. For example, as recognised by Fink et al. (2006), knowledge regarding the mathematical number line is thought to emerge from a "combination of an inherent understanding of how to estimate quantity and school-based instruction on the formal Hindu-Arabic number system" (Fink et al., 2006; p. 213). Based on the notion that the influence of biological factors may be more pronounced on tasks designed to access innate numerical competencies further research incorporating a more explicit measures of such skills would be particularly informative.

In summary the results of the present experiment revealed a significant positive correlation between right hand 2D:4D and number comparison scores in females aged 5-7 years old, suggesting a detrimental effect of PT on number comparison skills. These findings however were not replicated in males aged 5-7 years old or in either males or females aged 8-11 years old. The results also revealed a female advantage in counting performance in children aged 8-11 years old, the same advantage was not revealed for children aged 5-7 years old. Discrepancies in the findings across the two age groups may potentially imply the presence of social influences on performance. Ultimately however due to issues regarding the assessment of numerical skills clear conclusions and speculation regarding the findings of the present experiment are difficult and should be viewed with caution. In order to ensure that individual differences in basic mathematical skills are properly assessed future research may benefit from the use of reaction time measures. It would also be useful to more directly assess potential relationships between correlates of PT and more direct measures of innate numerical processing.

Chapter 3

Core numerical skills

3.1 Introduction

Experiment 1 attempted to replicate evidence from Fink et al. (2006) for a possible relationship between basic numerical skills and one potential correlate of prenatal testosterone (PT), namely the second to fourth digit ratio (2D:4D). Results revealed a significant positive correlation between 2D:4D and number comparison performance in females aged 5-7 years old, suggesting a detrimental effect of PT exposure on performance. The results of experiment 1 also revealed significant sex differences in performance on the counting task in children aged 8-11 years old (female advantage revealed). Such findings are generally in contrast to previous evidence which has suggested a lack of sex differences in basic numerical skills (Geary, 1996; Fink et al., 2006). Crucially however experiment 1 was subject to a number of methodological issues, including problems with the devised method of mathematic assessment and no control for the potential confound of ethnicity (see chapter 2), which should be considered when evaluating the findings.

As briefly highlighted in chapter one, mathematical performance is related to a variety of functionally and anatomically distinct cognitive components including those of a linguistic, spatial, and executive nature. Even basic numerical skills are likely to be related to a range of alternative cognitive processes. For example, a study by Kyttälä, Aunio, Lehto, Van Luit, & Hautamäki (2003) reported a significant relationship between counting performance and visuospatial working memory in children aged just 5-6 years old. In some instances, the cognitive components underlying performance may themselves be potentially influenced by exposure to PT (see chapter one for a review of evidence for a relationship between PT and cognition). In addition a variety of social and experiential factors may also exert an important influence on mathematical skills (see chapter 1). As a consequence, without fine-grained analysis of the underlying cognitive and/or social processes that may influence a particular task, a relationship between PT and performance on one mathematical task may offer very limited information as to the potential relationship between PT and performance on a different mathematical task. At a general level of interpretation therefore it is extremely difficult to deconstruct the nature of any revealed relationship between PT and mathematical ability and the various mechanisms by which a PT could be impacting upon numerical and mathematical competence.

As highlighted in chapter two, while general mathematical ability, even at a young age, is based on a combination of cognitive skills and influenced by a variety of social factors a growing body of evidence also suggests that the capacity to represent and manipulate numerical quantity may actually be a biologically-determined, innate, category-specific domain of knowledge, hard-wired in the brain and possibly shaped by evolution (Dehaene et al., 2003). This innate sense of numerical magnitude is typically divided into two systems, i.e., 1) an exact system, known as subitizing, for the representation of small quantities and, 2) an approximate system for the 'noisy' representation or comparison of large magnitudes. These two systems have been identified in both nonhuman primates and preverbal infants and are thought to form a foundation for the development and learning of higher mathematical skills (see Fiegenson, Dehaene, & Spelke, 2004 for review). While both experiment 1 and the study conducted by Fink et al. (2006) attempted to tap basic numerical skills, it is likely that performance on the tasks employed in both studies reflect a combination of innate numerical knowledge and school-based and experiential learning. An investigation of the relationship between correlates of PT and a more direct measure of these very simple, inherent numerical capabilities however would minimise the potential influence of alternative cognitive skills and/or social factors and may present one mechanism whereby PT may exert a general impact upon higher and developing mathematical ability.

Interestingly a study by Bull and Benson (2006) does present possible evidence for a potential association between innate numerical competencies and PT. These authors reported a significant association between 2D:4D and the so-called. 'Spatial Numerical Associations of Response Codes' (or 'SNARC') effect. The SNARC effect is a phenomenon in which smaller numbers are responded to faster with the left hand, and larger numbers with the right hand. As the SNARC effect is observed in tasks that do not require an assessment of numerosity, the effect is taken as evidence for automatic activation of basic magnitude, represented as a spatially organised analogue number line. Bull and Benson (2006) found that males and females with lower 2D:4D (higher PT exposure) displayed a significantly stronger SNARC effect as compared to those with higher 2D:4D (lower PT). Despite its hypothesised relationship with representations of basic magnitude however the SNARC effect remains heavily related to alternative cognitive processes. As well as an obvious visuo-spatial connection, the effect has also been shown to depend upon reading habits, to the extent that individuals accustomed to a right-left reading style may actually show a reverse effect (Dehaene, Bossini, Giraux, 1993). Thus, while the findings of Bull and Benson (2006) may reflect a relationship between an innate sense of numerical quantity and one potential correlate of PT further research utilising a more direct assessment of innate numerical processing is still required in order to confirm this possible association.

Subsequent chapters in the current thesis will attempt to more directly explore any possible relationship between 2D:4D, as a proxy of PT exposure and innate systems of numerical knowledge. The current chapter will firstly describe evidence for, and the nature of these core systems of magnitude representation in human adults, infants and animals. The neural basis of these core systems and evidence for their role in the expression of developing and higher numerical and mathematical skills will also be considered. The chapter will conclude by outlining the aims and hypothesis of the forthcoming chapter in the thesis.

3.2 Core system 1: Approximate representation of large quantities in adults

When human adults are presented with an array of objects under conditions that prevent counting they are able to determine a close approximation of quantity (see Dehaene, 1997; Feigenson et al., 2004; Hauser & Spelke, 2004). This ability is thought to reflect innate representations of approximate magnitude. While estimates under these conditions are rarely a precise identification of the number that corresponds to the quantity, they are invariably ‘in the neighbourhood’ in which the quantity lies. These approximate representations follow a distinct pattern whereby estimations of magnitude become increasingly less accurate as numerical size increases (size/magnitude effect) and the variability in performance increases linearly with the size of the number involved (Dehaene & Marques, 2002; Whalen, Gallistel, & Gelman, 1999).

When asked to compare the numerical magnitude of two separate arrays, without counting, approximate estimates can be predicted based on the ratio of the two numbers involved, i.e. the joint influence of absolute magnitude, and the numerical differences between the two values (called the ‘distance effect’). For example, 6 and 10 are just as easy/hard to discriminate as 12 and 20 and easier to discriminate than 12 and 16. This effect was demonstrated in a study by Barth, Kanwisher, & Spelke (2003) who showed that the approximate numerical comparison between large and small sets is similar when the ratio of the numerosities to be compared is the same. Similar to the discrimination or identification of other physical parameters therefore approximate numerical comparison and representation can be characterised by Weber's Law, which states that the change in stimulus intensity needed for an organism to detect a change is a constant of the original stimulus intensity rather than a constant amount (Brannon, 2006). The proportionality constant referred to is known as the as the Weber fraction.

Adults' ability to approximately identify or discriminate numerosities is identical for arrays of different modalities. Evidence shows that, in addition to object arrays, adults can perform numerical estimations and comparisons on sequences of actions (Whalen et al., 1999), sequences of sounds and light flashes and visuospatial arrays (Barth et al., 2003). Such studies demonstrate an identical ratio limit for arrays in different modalities. Adults' approximate discrimination of numerosity is also similar regardless of the presentation format, i.e. spatial vs. temporal, and just as accurate at comparing two quantities when the elements in the two arrays are presented in different modalities (auditory vs. visual) and formats (spatial vs. temporal), as when the elements in the two arrays are presented in the same modality and format (Barth et al., 2003). Finally, there is evidence that adults are able to use their approximate system of numerical representation to perform non-symbolic addition or subtraction on two arrays of either the same modality (i.e. two successive arrays of dots), or of different modalities (one array of dots and one sequence of sounds) when non-numerical variables are controlled (Barth et al., 2006).

3.3 Core system 1: Approximate representation of large magnitudes and animals and infants

Contributing to the notion that an approximate sense of numerosity may be innate, a growing body of recent evidence suggests that, like adults, pre-linguistic human infants can also form approximate numerical representations subject to the similar signature properties. Using a visual preference habituation technique Xu and Spelke (2000) explored 6 month old infants' abilities to approximately represent numerical magnitude. The infants were repeatedly shown images of either 8 or 16 dots until looking time substantially decreased. The infants were then tested with alternating arrays of 8 and 16 dots. Continuous variables such as display size, total filled area, and correlated properties including, surface brightness and texture, were carefully controlled for (see also Brannon, Abbott, & Lutz, 2004). The infants looked significantly longer at the numerically novel test array, regardless of whether they had been originally habituated to 8 or 16, suggesting sensitivity to the difference between the two numerosities that could not be accounted for by differences in non-numerical dimensions. Further experiments under identical conditions showed that infant were also able to discriminate arrays of 16 vs. 32 dots but failed to discriminate arrays of 8 vs. 12 or 16 vs. 24 (Xu & Spelke, 2000; Xu, Spelke, Goddard, 2005).

Similar findings have been revealed for arrays presented in different modalities (Lipton & Spelke, 2003), suggesting that infants possess an abstract, amodal, ratio-dependent ability to discriminate numerosity similar to that observed in adults under conditions that prevent counting. Infants are also able to form expectations about the outcome of arithmetic problems over large numerosities. For example, using a violation of expectation paradigm, infants were shown correct and incorrect outcomes to a mathematical operation (5+5 or 10-5), presented on a computer screen. With continuous variables controlled for, infants looked significantly longer at impossible outcomes, suggesting a basic understanding of addition and subtraction computations (McCrink & Wynn, 2004). Such findings have been reaffirmed in a study by Barth et al. (2005) in which children observed a computer screen as an array of blue dots appeared then disappeared behind an occluder; a second array of blue dots then appeared and disappeared behind the same occluder, and finally, an array of red dots appeared, and children were asked whether there were more blue or red dots. When tested with ratio differences of 0.57, 0.67 and 0.8 between the comparison and sum of the first two arrays, children performed significantly above chance.

Evidence that the ability to approximately represent numerical magnitudes may have evolved comes from similar evidence in animals to that revealed in infants. Operant conditioning studies show that animals trained to press a lever/ make contact with a key; *N* number of times or in response to *N* number of events or actions, produce a mean number of presses normally distributed around the target number, with a variability of errors proportional to the target number (see Brannon & Terrace, 2001; Hauser, 2000). These findings suggest that, like adults, animals also posses a ‘fuzzy’ representation of approximate numerosity, whereby accuracy of enumeration decreases with increasing magnitude.

Additional data suggests that, as well as the ability to represent approximate numerosity, a wide variety of animal taxa are also able to discriminate sets on the basis of approximate numerosity. Mirroring effects obtained in human adults and infants, this ability is governed by Weber's Law, such that performance for any given numerical discrimination depends on the ratio of the two numerosities being compared (Brannon & Terrace, 2002). A recent study by Cantlon and Brannon (2007) directly compared the performance of college students and two rhesus macaque monkeys on a number comparison task. As anticipated, findings revealed that for both groups, accuracy and latency were controlled by the ratio of the numerosities being compared across a wide range of values.

Like human adults and infants, animals' approximate representations of numerosity appear to be abstract and amodal, with evidence demonstrating that animals can represent and compare different types of entities, i.e. objects, tones, light flashes, and self-generated actions, across different styles of presentation, i.e. simultaneous or sequential (Church & Meck, 1984; Meck & Church, 1983). Further research has confirmed that patterns of performance are not due to other non-numerical factors such as total area, brightness, circumference, and density in simultaneous tasks, and inter-event interval and stimulus duration in sequential tasks (Emmerton 1998; Machado & Keen 2002; Meck & Church 1983; Nieder & Miller 2003; 2004).

While research on the numerical capabilities of animals has been criticised due to the fact that, unlike humans, animals require conditions of extensive training to perform these abilities, there is evidence that animals can spontaneously and automatically attend to the numerical attributes without training. For example, a study by Hauser et al. (2003) showed ratio dependent number discrimination in untrained tamarin monkeys. Further evidence from Brandon and Terrace (1998) illustrates that animals which have received training, posses the capacity to generalise to numerosities beyond the training range.

Finally, there is evidence that animals, like humans adults, can use their approximate representations of numerosity to compute addition operations over large sets. For example, in a study by Flombaum, Junge and Hauser (2005), rhesus monkeys observed 4 lemons on a stage and then watched as an occluder was placed in front of the lemons, blocking them from view. Four more lemons were then placed behind the occluder before it was removed to reveal a possible outcome of 8 lemons or an impossible outcome of 4 lemons. The monkeys looked significantly longer at the impossible 4 lemons outcome suggesting that their expectation had been violated. No significant differences in looking time however were revealed when the monkeys observed a 2 + 2 operation and tested with an outcome of 4 vs. 6 implying that this ability is; *a)* based on an approximate system of representation as opposed to an exact mechanism of computation and, *b)* subject to the same ratio dependence as number comparison such that accuracy, in this case, is dependent on the ratio between the observed and expected outcome.

3.4 Core system 2: Exact representation of small quantities in adults

Research suggests that human adults posses two distinct processes that may be utilised in the judgement of small numerosity. Firstly, a process named subitizing (Kaufman, Lord, Reese, & Wolkmann, 1949) characterised as the rapid and error free labelling of simultaneously presented small quantities, generally up to four, and secondly, the process of conventional counting. Evidence for this dichotomy is derived from reaction time data for the enumeration of visual arrays of dots. Latencies of enumeration mapped as a function of numerosity typically display a linear rise in response time with increasing numerosity (a slope of approximately 250-300ms/dot) for all integers greater that four. For quantities in the subitizing range however response times are faster with a far shallower incline (although see; Balakrishnan & Ashby, 1991; 1992 for contradictory evidence). Adults ability to rapidly identify quantities <4, or subitize, is hypothesised to reflect the second core system of numerical processing.

3.5 Core system 2: Exact representation of small quantities in animals and children

Similar to the first core system of numerical processing described, evidence suggests that the ability to subitize may present an innate, evolved numerical competency is derived from research suggesting that the ability to recognise and precisely enumerate small quantities is also present in animals and pre-linguistic infants. For example, utilising visual fixation techniques Antell and Keating (1983) habituated infants just 21-44 hours old to a fixed number of dots (between 1-4). The authors found that the infants subsequently gazed significantly longer at a slide containing a novel number of dots suggesting sensitivity to the numerical values of the display. Such findings have been demonstrated to be consistent over slides depicting sets of realistic objects of variable size, shape and spatial layout (Strauss & Curtis 1981; van Loosbroek

& Smitsman 1990). Similar studies have also demonstrated basic computational abilities in infants just six months of age (Sharron & Wynn 1998; Wynn 1996).

In one of the most widely cited empirical studies in the area Wynn (1992) reported evidence for the ability to track simple additive and subtractive numerical transformations in children as young as four to five months. In a 'violation of expectation paradigm' infants were seated facing a small stage on which they witnessed a physical transformation, for example two Mickey Mouse dolls sequentially placed behind a screen. The screen was then removed and the infants displayed surprise if the numerically appropriate number of objects was not observed. Such results have been replicated firstly by Simon, Hespos, & Rochat (1995) who, using characters from the TV show 'Sesame Street' further ruled out the possibility that biases in the identity of the dolls may account for Wynn's findings. A subsequent study by Koechlin (1997) also ruled out possible biases as a result of the location of the dolls. Similar findings have also been revealed using alternative methods. In a study by Feigenson, Carey and Spelke (2002) 10- and 12- month old infants watched as an experimenter placed different numbers of cookies into two separate boxes. Given the choice of 1 vs. 2 and 2 vs. 3 equal-sized crackers infants spontaneously and reliably crawled towards the box containing the greatest number of crackers. Given the choice of 3 vs. 4, 2 vs. 4, 3 vs. 6 and 1 vs. 4 cookies however, despite the highly discriminable ratio between the quantities, infants showed no preference in their approach patterns. Unlike approximate core representations of numerosity therefore the ability to represent and discriminate small numerosities does not appear to be governed by the numerical ratio but rather by the absolute number of items, with an upper-limit of 3. A similar 3 item limit was reported by Starkey and Cooper (1980) and Feigenson and Carey (2003).

Research also shows that infants are able to represent small numbers of visual-events and auditory sequences (e.g. puppet jumps and sounds). In a study by Bijeljac-Babic, Bertoncini and Mehler (1993) infants aged only four days old successfully discriminated two and three syllable words controlled for phonemic content, duration and speech rate. Evidence also exists for cross-modal numerosity matching across quantities < 4, for example in a study by Starkey et al. (1983) infants attended to two-object display longer when accompanied by two drumbeats than by three, and the three-object display was attended to longer when accompanied by three drumbeats than by two, suggesting infants are able to identify numerical correspondences across different kinds of stimulus modalities. Similar cross modal numerosity matching over small quantities was also revealed by Jordan and Brannon (2006) who found that 7-month-old infants looked preferentially at videos that contained the number of conspecifics that numerically matched the number of vocalisations they heard.

Infants' ability to represent small numerosities shows two important constraints. Firstly, a number of studies suggest that infants' ability to represent small numerosities may be confounded by continuous variables such as area (see Clearfield & Mix, 1999; Feigenson et al., 2002; Xu, 2003). Secondly, there is evidence that infants are unable to track small numbers of continuous substances or objects that appear and disappear discontinuously (Chiang & Wynn, 2000). Similar limitations have also been revealed on adults' ability to subitize (Hauser & Spelke, 2004; Scholl & Pylyshyn, 1999; van Marle & Scholl, 2003).

Research reporting evidence for an exact system of representation for small numerosities in a number of animal species has been conducted using similar methods to those employed with human infants. Using a violation of expectation, 'box search' technique similar to that employed by Feigenson and Carey (2003), Santos, Hauser, & Spelke (2001) found that untrained rhesus macaques searched longer in a box containing one object when expecting to find two objects than when expecting to find only one object, suggesting they can distinguish between one and two objects in the absence of training. A further example comes from Hauser, Carey, & Hauser (2000) who utilised a two box search method (employed in infants by Feigenson et al., 2002) to also explore small number discrimination in rhesus monkeys. Untrained rhesus monkeys watched as the experimenter sequentially placed apple slices into two separate boxes. Given choices of 1 vs. 2, 2 vs. 3 and 3 vs. 4 monkeys spontaneous approached the box with the largest quantity. Patterns of approach for quantities 3 vs. 8 and 4 vs. 8 however were at random. Similar to human infants therefore monkeys appear to possess a system for the representation of small numbers that is subject to a set size limit (despite being slightly higher than human infants), and independent of the ratio between the two numerosities. This finding remained similar after controlling for the possible confounds of volume and event timing.

Further evidence shows that rhesus monkeys are also capable of solving simple addition problems over small sets. Using a violation of expectation design similar to Wynn's (1992), replacing dolls for eggplants, monkeys looked longer at impossible than at possible outcomes in 1+1=1 vs. 2 comparisons, as well as in 2−1=2 vs. 1 comparisons. Similar findings were revealed in a subsequent study by Hauser and Carey (2003) in which rhesus monkeys succeeded in discriminating the correct outcomes of 1 + 1 (= 2 or 3), 2 + 1 (= 3 or 4), but failed with the computations 2 + 1 + 1 (= 3 , 4, or 5) and 1 + 1 +1 (= 2 or 3). The ability to discriminate 1 + 1 = 2 or 3 has also been revealed in laboratory reared cotton-top tamarins (Uller et al., 2001) and domesticated dogs (West & Young, 2002). A series of experiments conducted by Sulkowski and Hauser (2002) further confirmed rhesus monkeys ability to successful discriminate subtraction problems over small sets as well as addition problems.

3.6 Core systems of number and higher mathematical processing

As mentioned previously, the above described ontogenetically and phylogenetically developed core systems of numerical representation are hypothesised to constitute possible factors guiding the acquisition of formal arithmetic and higher mathematical processing. Both systems are thought to provide young children with a core representation of quantity and limited understanding of cardinality which when coupled with developing linguistic ability and ongoing mathematical experience/ education facilitate the acquisition of additional domains of numerical and mathematical competence, such as, a bourgeoning understanding and lexicon of number words, counting procedures, digits for written notation and procedures for calculation. It has been suggested that the process of subitizing for example, may facilitate the associations of specific quantities to their allied verbal and symbolic representations thus introducing semantic meaning to the first few number words and Arabic symbols (Gallistel & Gelmen, 1992; Wynn, 1990).

Support for the notion that core numerical skills may influence developing higher mathematical abilities can be derived from a recent study by Durand, Hulme, Larkin and Snowling (2005) that explored predictors of individual differences in arithmetic and reading in 7-10 year olds. The findings revealed that only digit comparison in addition to verbal ability was a significant independent predictor of arithmetic ability. The authors postulated that the digit comparison tasks tapped a basic aspect of children’s understanding of numerical magnitude, and that such basic magnitude representations may be one important source of developmental variations in arithmetic skills. Further evidence can be observed in a study by Benoit, Lehalle and Jouen (2004) which investigated the contribution of counting and subitizing processes to the acquisition of small-number words in 48 normally developing 3-5 year olds. Results showed that, for small quantities (<4) children were more likely to give the correct number-word when enumerating dot patterns presented simultaneously as compared to sequentially. The authors concluded that if the meaning of small-number words relies on counting, then performance would have been unaffected or even improved by sequential presentation, and thus suggest that the process of subitizing is more primitive than counting, and is likely to underlie the development of an understanding the first few number words.

The notion that these core systems of number may guide the acquisition of formal arithmetic and higher mathematical processing is further reinforced by evidence implying a ‘core deficit disorder’ in the expression of a subtype of developmental dyscalculia. A study by Llanderl et al. (2004) for example, explored the performance of children with dyscalculia, reading difficulties, or both on a range of basic numerical tasks in comparison to controls. Findings revealed deficits in speed of number comparison, dot counting and a trend towards deficits in subitizing despite normal performance on similar non-numerical tasks. The authors suggested that a lack of understanding of numerosity, and a poor capacity to recognize and discriminate small numerosities, may prevent dyscalculics developing the normal meanings for numerical expressions, and lead to their difficulties in learning and retaining information regarding numbers. The authors concluded that dyscalculia is the result of specific disabilities in basic numerical processing, rather than the consequences of deficits in other cognitive abilities (although see, Rousselle & Noel, 2007).

Research by Koontz and Berch (1996) also revealed that children with arithmetical difficulties demonstrated problems with subitizing, their finding suggesting that such children employ time-consuming counting strategies for set sizes as small as three items. Similarly, Geary, Hamson and Hoard (2000) found small but systematic group differences between first grade dyscalculic children and controls in magnitude comparison. Further evidence comes from a single-case study reported by Butterworth (1999) of a dyscalculic adult “Charles” who despite normal IQ and reasoning demonstrated deficits in number comparison and subitizing. In a similar case reported by Kaufmann (2002) a 14-year old boy showed no distance effect despite being able to complete multi-digit calculation procedures. Brain imaging evidence from Kucian et al. (2006) reported that children with mathematical difficulties exhibited weaker neural activation as compared to controls during approximate calculation, potentially implicating deficient recruitment of neural resources in the processing of numerical magnitudes.

A study by Desoete and Gregoire (2007) found deficits in subitizing and estimation of size to be features of delayed mental arithmetic and number knowledge ability in grade 1. Furthermore, in their sample of 20 males and 10 females, 33% of children with mathematical learning difficulties continued to struggle with subitizing at grade 3. Finally, there is evidence that girls with Turners syndrome, a congenital condition often associated with developmental dyscalculia in the context of normal general intelligence, demonstrate deficits in performance on number sense tasks, e.g. cognitive estimation, subitizing, addition and subtraction (e.g. Bruandet et al., 2004).

As adults, humans appear to retain these two core numerical systems and adopt them during quantitative reasoning tasks. This can be seen in research where adults have been required to compare the relative magnitudes represented by two Arabic numerals. Findings show that reaction times are influenced by distance and size effects (Moyer & Landauer, 1967), thus, although number can be represented with arbitrary symbols (e.g. 2 and two) a non-verbal representation with an analogue format appears to underlie these symbols.

3.7 Neural basis of core systems of number

An expanding body of literature has attempted to identify the neural mechanisms that may underlie these two core systems of number. The potential neuroanatomy of the approximate number system in particular has received considerable attention and is fairly well documented (Dehaene, 1997; Dehaene et al., 2003; Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999).

Initially discovered on the basis of lesion data (Gerstmann 1940; Henschen 1919; Roland & Friberg 1985) and later replicated using Positron Emission Tomography: PET (Dehaene et al., 1996; Pesenti, Thioux, Seron, & De Volder, 2000) and functional Magnetic Resonance Imaging: fMRI (Burbaud et al., 1999; Rueckert et al., 1996) the involvement of the parietal cortex, specifically the inferior parietal cortex, in tasks of number processing and mental calculation is well established. It has been consistently found that magnitude comparison (Dehaene et al., 2003; Dehaene et al., 1999; Temple & Posner, 1998), mental number line (Zorzi et al., 2002), and many arithmetic tasks (Chochon, Cohen, van der Moortele, & Dehaene, 1999; Rivera, Reiss, Eckert, & Menon., 2005) engage the bilateral intraparietal sulcus, although other regions are also engaged e.g., frontal regions associated with working memory (Rivera et al., 2005).

Based on a synthesis of existing literature Dehaene et al. (2003) postulated a functional characterisation of the neuroanatomical loci for number processing in the parietal lobe based on the existence of three distinct circuits co-existing to subserve different aspect of number processing ('The Triple Code Model'). Firstly, a region of the left angular gyrus associated with verbal processing of number, secondly the posterior superior parietal lobule associated with visuospatial processing and the allocation of spatial and non spatial attention, and finally of specific interest to the current review, the horizontal bilateral segment of the intraparietal sulcus (referred to as HIPS) associated with the core representation of approximate quantity. A number of lines of evidence support the suggestion that the HIPS constitutes the site of our apparently core ability to approximately represent numerical magnitude.

Firstly, the HIPS is more active when estimating the approximate result of an addition problem than when computing its exact solution (Dehaene et al., 1999). When carrying out exact calculations the HIPS is also shows greater activation for subtraction as compared to multiplication (Chochon et al., 1999; Lee, 2000). As multiplication tables and small exact addition facts are typically stored in rote verbal memory whereas subtraction problems are generally not stored in verbal memory, Dehaene et al. (2003) suggests that this may reflect evidence that the HIPS is more active for numerical operations that require a genuine manipulation of quantity.

Secondly, the HIPS is more active when an arithmetic operation requires a quantitative representation of numbers, for example the HIPS is more active when participants engage in calculation that when they merely have to read numerical symbols (Burbaud et al., 1999; Chochon et al., 1999; Pesenti et al., 2000). Similarly, there is evidence that the HIPS is also more active when comparing the magnitude of two numbers than when simply reading them.

Thirdly, the HIPS shows a robust category specificity for numbers and number processing, when directly contrasted with different categories or objects, including those which can be characterized along non numerical scales (e.g. the ferocity of animals, relative position of body parts, the orientation of two visually presented characters). Even on simple detection tasks that do not require numerical judgement the HIPS is the only region that shows higher activation when processing numbers then when processing letters of the alphabet or colours (Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003) despite that fact that the alphabet shares with numbers a strong serial component; and both letters and colours, (like numbers) show a distance effect (e.g., when detecting the letter D it takes participants longer to reject the letter C than the letter N).

Fourth, when attending to numerosity, the HIPS is similarly activated for Arabic numerals, number words, and non-symbolic stimuli presented in different modalities, e.g. dots, tones. Similar to the signature properties of core representations of approximate numerosity therefore the region appears to possess an abstract representation of number, independent of the modality in which the numerical stimulus is presented.

Fifth, also in line with the behavioural signature properties of core approximate representations of number, activation of the HIPS is modulated by the absolute magnitude of the number/s and, in a number comparison task, the numerical distance between the two magnitudes. HIPS activity lasts longer with small numbers than with large numbers (Kiefer & Dehaene, 1997; Stanescu-Cosson et al., 2000) and is reduced as the numerical distance between two numbers being compared decreases (Pinel et al., 2001).

Sixth, lesion data suggests the existence of distinct semantic systems for numerical quantity representations and operations based in the intraparietal sulcus. Evidence from single-case studies furthermore suggests that the representation of number doubly dissociates from other categories of words at the semantic level. Single-case studies in individuals with lesions broadly affecting the left temporo-frontal cortices, but sparing intraparietal regions have been reported to show spared calculation and number comprehension abilities in spite of gross deficits in semantic processing or semantic dementia (Butterworth, Cappelletti, & Kopelman, 2001; Cappelletti, Butterworth, & Kopelman, 2001; Thioux et al., 1998). Further evidence from single-case studies shows that an understanding of numbers and their relations can be specifically impaired in the context of preserved language and semantics (Cipolotti, Butterworth, & Denes, 1991; Dehaene & Cohen, 1997; Delazer & Benke, 1997), with the majority of these cases resulting from lesions in the parietal regions, particularly in the left hemisphere.

There is evidence for a particular sub-category of patients who appear to suffer from a category-specific impairment of the numerical magnitude system. Such individuals can comprehend and produce numbers in all formats but are unable to subtract, compare or bisect numbers regardless of the format in which the number is presented. One patient known as ‘MAR’ (Dehaene & Cohen, 1997) displayed impairments in deciding which of two numerical values was larger, and was almost entirely unable to identify which number falls in the middle of two other numbers (numerical bisection), but could easily perform analogous comparison and bisection tasks in other non-numerical domains, e.g. for days of the week, months, or letters of the alphabet. Such deficits appear to be specific to the number domain, but not due to the inability to identify numbers or produce the operation result.

Further developmental evidence suggests that early deficits in the HIPS system may correlate with developmental dyscalculia. A study by Levy, Reis and Grafman (1999) for example, reported the case of an individual with lifelong dyscalculia, but with superior intelligence and reading ability. While MRI data appeared normal in this individual, magnetic resonance spectroscopy revealed a metabolic abnormality in the left inferior parietal area. Reduced gray matter in the left IPS associated with developmental dyscalculia has also been found in adolescents born at equally severe stages of prematurity (Issacs et al., 2001). Similar research has demonstrated a region of reduced grey matter and reduced activation in the right HIPS in females affected by Turner's syndrome (a condition in which mathematical learning difficulties are consistently reported).

Finally, evidence for the critical role of interparietal regions in the representation of approximate magnitude comes from recent research demonstrating evidence for the existence of specific IPS neurons tuned to approximate numerosity in monkeys. Research shows that the parietal sulcus is also active when non-human animals engage in numerical activities (Sawamura, Shima, & Tanji, 2002). Using a numerosity matching task a series of studies by Nieder and colleagues have revealed the existence of neurons in the depth of the IPS which fired selectively when a number of dots was presented visually regardless of non-numerical stimulus attributes such as circumference surface area or density. These neurons had identically short firing rates for all of the numerosities tested suggesting parallel extraction of numerosity across the entire display (e.g. Neider & Miller, 2003; 2004). The firing latencies for these IPS neurons were also revealed to be shorter than those typically observed for prefrontal neurons in response to number implicating that numerical information is first extracted and represented in the IPS before being transmitted to prefrontal circuits as needed for the requested task (Dehaene & Changeux, 1993). Crucially, individual neurons showed a maximal firing rate to one numerosity, and decreased firing rate as a function of distance from the preferred numerosity. Such representations therefore are approximate, and, similar to evidence at the behavioural level, subject to Weber's Law.

In addition to evidence regarding the localization of our core system for approximate numerical representation, research also suggest a certain degree of functional lateralization for numerical information in the parietal cortex which may have implications for core representations of number. Parietal activation appears greatest in the right hemisphere during some aspects of mental arithmetic (Mennon et al., 2000), and number comparison (Chochon et al., 1999; Stanescu-Cosson et al., 2000), whereas left frontal, angular gyrus, and cingulate cortices are strongly activated during the retrieval of exact arithmetic facts (Dehaene et al., 1999). Regions within the two hemispheres thus appear to be differentially engaged for different quantitative abilities, with a right-hemisphere advantage for tasks requiring more abstract (e.g., relative magnitude) numerical relations and a left-hemisphere advantage for tasks requiring more discrete quantitative information (e.g., Langdon & Warrington, 1997).

In comparison to evidence exploring the neural correlates of core representations of approximate numerosity, far less research has attempted to identify the neural basis of our core system for the exact representation of numerosities < 4. Feigenson et al. (2004) suggested that this might be due to the fact that subitizing is a basic, automatic function of early extrastriate areas, and that the representation of distinct objects is so fundamental to perception and cognition that it might elude current neuroimaging methods, where a control task must be devised in which the target system is not activated. The evidence that does exist has focused specifically on the process of hemispheric specialization during subitizing tasks. Findings here however are mixed. While a number of authors attribute the process of subitizing to the right hemisphere (e.g. Jackson & Coney, 2004; Kimura, 1996; Pasini & Tessari, 2001) Butterworth (1999) suggested that subitizing may actually be dominated by the left hemisphere. With regard to brain imaging evidence however, a lack of hemispheric domination for subitizing has been reported (Piazza, Mechelli, Butterworth, & Price, 2002; Sathian et al., 1999).

3.8 Aims of the forthcoming chapters

Given the evidence described above, a greater understanding of any potential mechanisms which may impact upon core numerical processes could have important implications to our understanding of individual differences in numerical and mathematical skills. The following chapters will attempt to explore any potential relationship between 2D:4D, as a potential marker of PT exposure, and performance on tasks designed to directly tap into these core numerical skills. If, as evidence suggests, our intuitive understanding of quantity and magnitude provides a foundation for the development and learning of higher mathematical and numerical skills then any potential relationship between the two may present one possible mechanism, by which PT might relate to mathematical abilities in a way that is unlikely to be secondary to alternative cognitive skills or social factors.

This is to not suggests that a link between PT and core numerical skills is the only avenue by which PT may impact upon mathematical ability, or that PT is the only or most important factor in the development of core numerical skills or higher numerical and mathematical skills, simply that, exposure to PT may present one variable influencing core numerical processes and thus potentially, developing and higher numerical and mathematical ability. Based on the findings of Bull and Benson (2006) it is hypothesized that lower (more masculinised) 2D:4D ratios will relate to better performance on tasks designed to assess core numerical skills. Given the revealed inconsistencies in previous literature regarding the presence of significant effects depending upon sex, the possibility that the strength and, perhaps, the nature of associations may differ between males and females will also be considered.

Chapter 4

Experiment 2: 2D:4D and Basic Processes of Enumeration in Adults.

4.1 Introduction

As detailed in the previous chapter, subitizing is defined as the rapid and error-free labelling of simultaneously presented small quantities, generally up to four items (Kaufman et al., 1949). Evidence suggests that the process of subitizing is present in infants and animals, and independent of higher-level cognitive processing, such as language (Strauss & Curtis, 1981; Antell & Keatin, 1983; van Loosbroek & Smitsman, 1990; Hauser, 2000; Feigenson & Carey, 2003, see chapter three for review) leading to the hypothesis that subitizing may constitute one aspect of innate, ‘core’ numerical processing (Feigenson et al., 2004).

In contrast to subitizing, the precise representation of larger quantities (>4) is thought to rely on a process of conventional counting. Counting can be characterised as process by which each successive quantity under enumeration represents an augmentation of the preceding number in the sequence by one, resulting in incremental increases in reaction times of enumeration as a function of increasing quantity (Halpern et al., 2007). While both subitizing and counting constitute mechanisms of precise enumeration, the two processes are considered to be distinct. Traditionally, the dichotomy between the two has been based on evidence for a sharp discontinuity in quantification performance whereby a shallow, non-linear sloped reaction time (RT) function is observed for the recognition of quantities up to approximately four, and a steeped sloped linear RT function is observed for the precise enumeration of cardinalities greater than four (Chi & Klahr, 1975; Mandler & Shebo, 1982; Svenson & Sjoberg, 1983; Trick & Pylyshyn, 1993, 1994, although see Balakrishnan & Ashby, 1991; 1992).

More recently, additional support for the dichotomy can be derived from functional imaging studies where a sudden increase of activity in posterior parietal, occipital and frontal regions for the enumeration of integers within the counting as compared to subitizing range has been reported (Nan et al., 2005; Piazza et al., 2002; 2003; Sathian et al., 1999). Further support for the distinction between subitizing and counting can also be deuced from neurological research in which apparent dissociations between the two processes have been observed (Dehaene & Cohen, 1994; Butterworth, 1999; Cipolotti et al., 1991). Developmentally, in line with the hypothesis that subitizing may constitute one aspect of core numerical processing, subitizing is thought to be a more primitive process of enumeration with evidence suggesting that the process of subitizing may contribute to the acquisition of developing counting skills (see chapter 3). The processes of counting and subitizing therefore provide a unique opportunity to explore possible relationships between 2D:4D and one aspect of 'core', innate numerical processing in contrast to a similar yet distinct basic numerical skill.

Intriguingly, there is also behavioural evidence that the two processes may demonstrate different patterns of hemispheric specialization. In a series of RT experiments Pesini and Tessari (2001) provided evidence for a left visual field (right hemisphere) advantage in the identification and comparison of quantities in the subitizing range, and a right visual field (left hemisphere) advantage during a comparison task for quantities in the counting range. Such evidence is of important theoretical interest when considering potential prenatal testosterone (PT) effects, given speculation that the hormone may play a significant role in the development of patterns of lateralization. In particular, Geschwind and Galaburda's (1987) hypothesis that the fetal testosterone may slow normal development of areas in the left hemisphere, leading to compensatory growth in corresponding regions of the right hemisphere, would imply that any potential influence of PT may exert opposite effects on the two enumeration process of counting and subitizing.

The current study aimed to investigate any potential associations that may exist between 2D:4D and subitizing and counting performance. Based on the findings of Bull and Benson (2006), it is hypothesized that lower (more masculinised) 2D:4D ratios will relate to increased automaticity for the primitive process of subitizing, and thus, decreased reaction time during the enumeration of quantities in the subitizing range. In line with previous evidence for a possible relationship between correlates of PT and counting (Finegan et al., 1992; Fink et al., 2006) it is also predicted that an association may be found between 2D:4D and enumeration of quantities in the counting range, given however the existing contradictions in prior evidence no directional predictions are formed regarding the nature of this possible relationship. Based on previous literature it is also hypothesised that different patterns may emerge in the strength, and perhaps, nature of associations between 2D:4D and subitizing and counting between males and females.

The present study also aimed to explore any potential relationships between PT and lateralization for the process of subitizing in contrast to the process of counting by examining reaction time and accuracy for enumeration of quantities in the subitizing (1-4) and counting (6-8) ranges for stimuli presented to the left and right visual fields. Based on evidence from Pesini and Tessari (2001) it is anticipated that different patterns of lateralization may be revealed for the two separate enumeration processes, such that a right hemisphere bias might be observed during subitizing, and a left hemisphere bias be seen during counting. Of greater interest to the current thesis, if, as according to Geschwind and Galaburda (1987), high fetal testosterone is associated with facilitated right hemispheric growth, it could be predicted that prenatal testosterone, and in turn 2D:4D measures, may demonstrate differential relationships with subitizing as compared to counting. As well as any potential simple correlations between 2D:4D and numerical performance therefore it is the potential interactions between the factors of 2D:4D, sex, visual field and task that are of particular interest. In order to facilitate mulitfactorial analysis and thus consideration of these potential complex interactions raw 2D:4D data in the present experiment was also used to categorise participants into low vs. high 2D:4D groups.

Building upon the limitations of the experiment one, the current study recorded and controlled for the potential influence of ethnicity on 2D:4D measures (see chapter 2). In addition, by adopting reaction time measures, individual differences in performance can still be evaluated regardless of any potential ceiling effects such as those indentified in experiment 1.

In an effort to control for the possibility that any associations between subitizing and/or counting and 2D:4D could potentially be secondary to a more general relationship between PT and RT/accuracy to any numerical stimuli and/or generic performance on any speeded response task, two control tasks are included in the procedure. One is aimed at the assessment of generic RT/accuracy on a non numerical task, the other is aimed at the evaluation of generic RT/accuracy on a numerical task.

4.2 Method

4.2.1 Design

Basic associations between 2D:4D subitizing and counting task performance were explored using correlations. In order to also consider any possible complex interactions between 2D:4D, sex, performance on the two separate numerical tasks and lateralization for numerical performance , a 2 x 2 x 2 x 2 quasi-experimental, mixed measures factorial design was employed including the factors, 2D:4D (high vs. low – determined via median split), sex (males vs. females), visual field (right vs. left) and process (subitizing vs. counting). Measures of general reaction time were available to be employed as covariates where required.

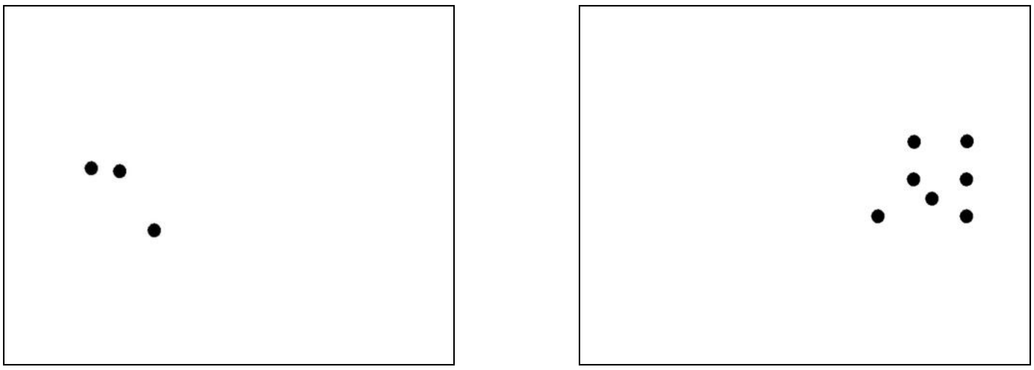
4.2.2 Participants

Eighty-nine participants were recruited from the undergraduate and postgraduate population of Northumbria University to take part in the experiment. All participants gave their full written informed consent to partake in the study and completed a brief biographical questionnaire. All participants reported normal or corrected-to-normal visual acuity and no participants reported any hormonal abnormalities or the consumption of any hormone-influencing drugs (excluding the contraceptive pill). Any individuals reporting injury to the second or fourth fingers of either hand were removed from the data (n = 7). In line with the criteria adopted in experiment 1 all left handed participants (assessed according to self-reported writing hand by means of a single biographical question) were also removed from the analysis (n=1). As 2D:4D is known to vary with ethnicity (Manning, Barley, et al., 2000; Manning, 2002) and may also be associated with an individual's sexual orientation (Putz et al., 2004; Rahman & Wilson, 2003; McFadden & Shubel, 2002; Robinson & Manning, 2000), the self identified ethnic origin and sexual orientation of each participant was also recorded. As only a small minority of the sample were not heterosexual (n = 4) and of white British origin (n = 1), in order to control variation according to ethnicity and sexual orientation, analysis employed only participants of the majority group for each dimension. In total, data recruited from 46 females (mean age – 21.04 years; SD – 5.6) and 30 males (mean age – 22.43 years; SD – 5.04) were included in the analysis. On completion of the study each participant received a payment of £4.

4.2.3 Measures

4.2.3.1 Reaction time tasks

RT tasks were administered using a standard PC with a 17-inch monitor, through a custom-made programme using Visual Basic. Visual stimuli for the RT tasks were developed using Microsoft 'Paint'. Experimental stimuli consisted of the quantities 2, 3, 4, 6, 7 or 8, presented in the form of black dots (0.8º visual angle) on a white background, to the left or right visual fields. All dots patterns were arranged within an invisible parameter of a 7.5×7.5 cm square with the internal side of each square positioned precisely 4cm from a central fixation, so that each dot pattern was viewed at an eccentricity varying from 4˚ to 11.5˚ of visual angle. Three sets of different dot patterns were randomly generated for each quantity and presented in a pseudo-random order so that each quantity appeared a total of 30 times (15 to the right, 15 to the left). The dot patterns were created using a random number generator in order to place elements into one of 36 numbered grid sections within the 7.5×7.5cm square. Maximum and minimum distances between items were thus 0.45˚ and 5.45˚, respectively, in the horizontal and vertical directions.



a) **b)**

Figure 2. Examples of; a) subitizing stimuli presented to the left visual field and b) counting stimuli presented to the right visual field.

Prior to every trial a central fixation point (X) appeared on the screen for a duration of 1000ms, followed after a 50ms pause by the stimuli. Participants were required to identify the number of dots presented on the screen. Stimuli remained on the screen until participant response. Following participant response, the inter-trial interval was randomly set between 1000ms and 1500ms. The task began with a practice block of 10 trials, 5 leftward presentations, 5 rightward presentations, of randomly selected arrays.

As explained above, two control tasks were also included in order to assess generic performance on speeded response tasks, namely; a control Arabic task, in which participants were asked to identify the Arabic digit (always 2, 3, 4, 6, 7 or 8) presented to the centre of the screen and finally a control letters task, in which participants were asked to identify whether a centrally presented letter was a vowel or a consonant (possible stimuli consisted of vowels; A, E, I, O, U and randomly selected consonants; B, L, N, S, Y). Target digits and letters on the control tasks were presented in black 72 point Times New Roman font.

Both tasks consisted of 10 randomly presented trials proceeded by a practice block of 5 trials, using similar parameters of presentation and duration to the experimental task. Responses were measured to the nearest millisecond via a 6 button response box. For experimental, and control Arabic task trials, numbers on the response box were presented linearly, in numerical order - numbers 2, 3, 4 were responded to using the left hand index finger while numbers 6, 7, 8 were responded to with the right hand index finger. Responses to the control letters tasks was made using the two most proximal keys on the six button response box, vowels were responded to using the right key, with the right hand and consonants by pressing the left key with the left hand. All participants initially completed the control vowel-consonant task, followed by the control Arabic task and finally the experimental subitizing/counting task.

4.2.3.2 Second to fourth finger ratio measures

The procedure adopted in order to calculate 2D:4D and evaluate the reliability of 2D:4D measures was identical to that adopted in experiment 1 (see chapter 2). Similar to experiment 1 intraclass correlation coefficients (r_I) showed high retest-reliability between first and second measurements of both the right (second r_I = 0.971; fourth r_I = 0.986; 2D:4D r_I = 0.813) and left hands (second r_I = 0.984; fourth r_I = 0.984; 2D:4D r_I = 0.87). For second digit measurements TEM was computed to be 0.87 and 0.677 with the rTEM calculated to be 1.224% and 0.948% for the right and left hands respectively. TEM for fourth digit measurements was calculated at 0.684 for right hand measures and 0.677 for left hand measures with corresponding rTEM calculated to be 0.948% and

1.007% respectively. In line with experiment 1 therefore an acceptable degree of precision for second and fourth finger measurements was met (see Weinberg et al. 2005).

As anticipated, 2D:4D ratios were lower in males (right hand mean = 0.977, SD = 0.025; left hand mean = 0.981, SD = 0.026) than in females (right hand mean = 0.991, SD = 0.031; left hand mean = 0.989, SD = 0.032) sex differences however were only significant for right hand ratios, (right hand 2D:4D - $t_{(74)} = 2.106$; $p = 0.039$, left hand 2D:4D – $t_{(74)} = 1.181$; $p = 0.241$.

Within-sex median splits according to 2D:4D for both the right (median males = 0.976, median females = 0.987) and left (median males = 0.978, median females = 0.982) hands were applied to the data. On the basis of this median spit two sets of high vs. low 2D:4D groups were formed. Mean 2D:4D values were confirmed to be significantly different between these groups, see Appendix 2 for t-test results.

4.2.4 Procedure

In order to restrict head movement during the computerised task all participants were requested to place their head on a chin rest positioned approximately 50cm from the centre of the monitor and focus their gaze towards the centre of the screen. Participants completed the computerised experimental and control task and RTs and errors were recorded. Participants were instructed to respond as fast as they could while also paying attention to accuracy. Digit measurements were taken twice, once prior to conducting the experiment and once following completion. The entire experimental procedure lasted approximately 30 minutes. All participants were fully debriefed. The protocol was approved by Northumbria University, School of Psychology and Sport Sciences Ethics Committee.

4.3 Results

Full tables of means for 2D:4D measures, subitizing and counting performance in both males and females separately and the sample overall can be viewed in Appendix 2.

Similar to experiment 1 data was assessed for normality using one-sample Kolmogorov-Smirnov tests. Analysis of normality revealed that subitizing reaction times and both subitizing and counting percentage error scores were not normally distributed, see Appendix 2. Where relevant therefore non-parametric tests were adopted in order to explore behavioural aspects of the data and simple correlations between 2D:4D and performance. As highlighted in the introduction however one of the primary points of interest in the current study is the potential complex interactions that might exist between 2D:4D, sex and subitizing performance relative to counting performance for information presented to the left and right visual fields. Given that there is no non-parametric method by which complex interactions between all four factors may be considered and that any non-parametric one-way analysis would offer no further information to that which can be derived via simple correlational analysis, 4-Way ANOVA analysis was adopted in order to explore any main or interaction effects of digit ratio (low vs. high; separate ANOVAs conducted for groups formed on the basis of either right or left hand measures), sex (male vs. female), process (subitizing vs. counting), and visual field (right vs. left) and their possible interactions on performance. The results of these ANOVA statistics however should be considered in light of the relative loss of test efficiency that may exist due to violations of normality.

4.3.1Behavioural reaction time data

Means and standard deviations for reaction times to correct responses and percentage error for performance on computerised tasks were calculated. Figure 2 and tables 5 and 7 illustrate mean RTs and percentage error scores on the experimental tasks and mean increases in percentage error and RTs as a function of increasing numerosity. The Wilcoxn signed ranks test was used to explore differences in subitizing reaction time and subitizing and counting percentage error with increasing numerosity. T-tests were used to explore changes in counting reaction times as a function of numerosity. Bonferroni corrections for multiple comparisons were applied when considering differences between numerosities ($\alpha = 0.0125$). Both the descriptive and inferential statistics illustrated in figure 2 and table 5 suggest that patterns of RT during the experimental task did not match the anticipated response characteristics of the two distinct enumeration processes. Reaction times to quantities in the subitizing range are typically reported in the range of approximately 500-800ms with a RT x Set size slope function of approximately 40–100ms per item (Trick & Pylyshyn, 1994) typically identified, although a number of authors do report a lower slope gradient of approximately 20-30ms per item (e.g. Wender & Rothkegel, 2000). Dehaene and Cohen (1994) however offer a more comprehensive description of the subitizing RT x set slope reporting an increase of at most 20ms from 1 to 2 items, approximately 50ms from 2-3 items, and 100-200ms from 3-4 items.

Generally therefore, participants from the current study were slightly beyond the range of usual reaction times, and showed a steeper incline than that typically reported for quantities 2-3. With regard to reaction times to quantities 6-8, the process of counting is typically reported to yield a RT x Set size slope of approximately 250 – 350ms per item. While this pattern was revealed in responses to 6 vs. 7 items, RTs to patterns containing 8 dots were actually lower than to those containing 7 dots. Given the corresponding increase in percentage error rate for responses to 7 dot arrays in comparison to 8 dot arrays, this pattern is more characteristic of an estimation strategy where accuracy and RT is reported to be lower than that observed for the process of counting (Kaufman et al., 1949; Mandler & Shebo, 1987). All future analysis for the process of counting in the current study therefore included averaged data for the quantities 6 and 7 only.

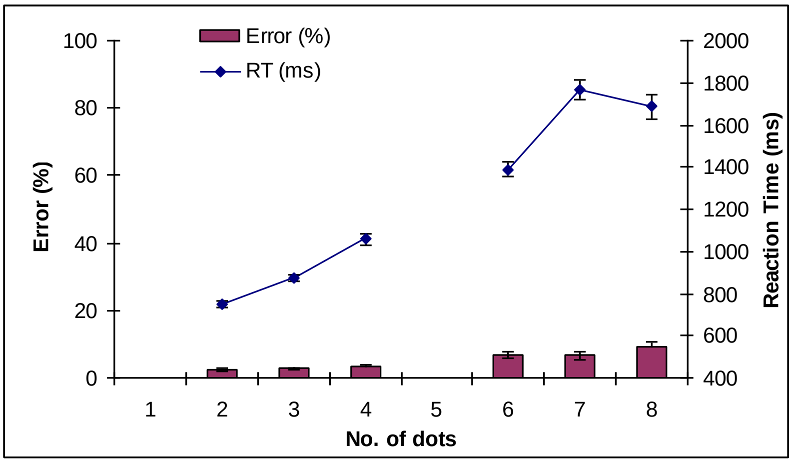


Figure 3. Mean response times and average percentage error for the enumeration of arrays in a subitizing (2-4) and counting range (6-8), including error bars indicating SEM.

Table 5

Average values and analysis of increases in RT and percentage error as a function of ascending numerosity (n = 76), significant values indicated in bold.

		RT			Percentage error		
	Quantity	Mean increase (SD)	Statistic	p	Mean increase (SD)	Statistic	p
Subitizing range	2-3	121 (90)	**Z = -7.031**	**<0.001**	0.3 (4.39)	Z = -0.508	0.612
	3-4	185 (153)	**Z = -7.574**	**<0.001**	0.4 (4.8)	Z = -0.741	0.459
Counting range	6-7	378 (237)	**t = 13.936**	**<0.001**	0 (9.61)	Z = -0.289	0.773
	7-8	80 (274)	**t = 2.542**	**0.013**	2.54 (10.88)	Z = -1.757	0.079

Reaction times to the enumeration of quantities in both the subitizing and counting ranges decreased with increasing percentage error. As can be seen in table 6 significant speed-accuracy associations were revealed for responses to quantities in the subitizing range overall (averaged from left and right visual field presentations), and for quantities in the subitizing range specifically presented to the right visual field. For quantities in the counting range significant speed accuracy associations were identified for information presented to the left visual field and overall (averaged from left and right visual field presentations).

Table 6

Spearman's correlation (ρ) analysis of speed accuracy associations for performance on the RT task (n = 76).

	Left	**Right**	**Overall**
Subitizing	ρ = -0.204, p = 0.077	ρ = -0.281, p = 0.014	ρ = -0.322, p = 0.005
Counting	ρ = -0.422, p < 0.001	ρ = -0.18, p = 0.121	ρ = -0.338, p = 0.003

Notably however, percentage error scores were generally very low (see table 7), with rates of percentage error not exceeding 7% during either subitizing or counting, all further investigations of a potential effect of 2D:4D in the current chapter therefore focused exclusively on RT data.

Table 7

Means and standard deviations for participant (n = 76) RTs in ms to correct responses and percentage error for the enumeration processes of subitizing and counting, overall and for information presented to both the right (RVF) left (LVF) visual fields.

	RT (ms) - Mean(SD)			***Percentage Error - Mean(SD)***		
Task	*LVF Response*	*RVF Response*	*Overall*	*LVF Response*	*RVF Response*	*Overall*
Subitizing	879 (144)	904 (150)	891 (146)	2.6 (3.4)	3.36 (3.13)	2.98 (2.81)
Counting	1579 (336)	1575 (326)	1577 (327)	6.75 (7.44)	6.93 (7.05)	6.84 (6.6)

4.3.2 Correlations

Table 8 shows the results of Spearman's correlation analysis to explore associations between 2D:4D and subitizing reaction times. Table 9 displays the results of Pearson's correlation analysis to explore associations between 2D:4D and counting reaction times.

Power for each analysis was calculated using G*Power (Faul et al., 2009). Power for the Spearman's analysis was computed based on adjusted sample sizes in order to take into account the relative loss of power for a Spearman's analyses relative to a Pearson's using the method described in experiment 1 (see section 2.3.1). Power calculation again showed that power was low for all correlation analyses. Small to moderate effect sizes were observed for a number of the correlation coefficients, i.e. between left hand 2D:4D and both subitizing and counting performance in males and right and left hand 2D:4D and subitizing in females and right hand 2D:4D and counting performance in females. Where such effects were observed positive correlations were revealed in males, suggesting slower subitizing and counting reaction times in higher 2D:4D participants, while negative correlations were revealed in females, conversely suggesting faster reaction times in higher 2D:4D participants. None of the associations between 2D:4D and either subitizing or counting reaction times in overall data or in male and female data analysed separately however were found to be significant.

Table 8

Spearman correlation coefficients (ρ) for the relationship between left and right hand 2D:4D and subitizing performance. P values and power calculations (1-β) for each analysis are also listed.

	Males and Females n = 76		**Males n = 30**		**Females n = 46**	
	Right 2D:4D	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Subitizing Overall	ρ = -0.034 p = 0.771 1-β = 0.059	ρ = -0.014 p = 0.904 1-β = 0.052	ρ = 0.093 p = 0.625 1-β = 0.075	ρ = 0.24 p = 0.201 1-β = 0.235	ρ = -0.161 p = 0.285 1-β = 0.175	ρ = -0.202 p = 0.178 1-β = 0.251
Subitizing Left Visual Field	ρ = -0.058 p = 0.621 1-β =0.076	ρ = -0.031 p = 0.791 1-β = 0.057	ρ = 0.043 p = 0.82 1-β = 0.055	ρ = 0.213 p = 0.258 1-β = 0.194	ρ = -0.158 p = 0.295 1-β = 0.17	ρ = -0.201 p = 0.249 1-β = 0.181
Subitizing Right Visual Field	ρ = 0.013 p = 0.91 1-β =0.051	ρ = 0.036 p = 0.756 1-β = 0.06	ρ = 0.139 p = 0.463 1-β = 0.108	ρ = 0.282 p = 0.131 1-β = 0.311	ρ = -0.108 p = 0.477 1-β =0.104	ρ = -0.137 p = 0.364 1-β = 0.139

Table 9

Pearson's correlation coefficients (r) *for the relationship between left and right hand 2D:4D and counting performance. P values and power calculations (1-* β) *for each analysis are also listed.*

	Males and Females n = 76		**Males n = 30**		**Females n = 46**	
	Right 2D:4D	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Counting Overall	r = -0.019 p = 0.871 1-β = 0.053	r = 0.064 p = 0.583 1-β = 0.086	r = 0.047 p = 0.807 1-β = 0.057	r = 0.148 p = 0.434 1-β = 0.124	r = -0.143 p = 0.342 1-β = 0.16	r = -0.033 p = 0.826 1-β = 0.056
Counting Left Visual Field	r = -0.008 p = 0.948 1-β = 0.05	r = 0.094 p = 0.422 1-β = 0.128	r = 0.074 p = 0.697 1-β = 0.068	r = 0.172 p = 0.365 1-β = 0.152	r = -0.136 p = 0.367 1-β = 0.149	r = 0.006 p = 0.969 1-β = 0.05
Counting Right Visual Field	r = -0.03 p = 0.797 1-β = 0.058	r = 0.032 p = 0.094 1-β = 0.059	r = 0.017 p = 0.93 1-β = 0.051	r = 0.121 p = 0.524 1-β = 0.1	r = -0.146 p = 0.332 1-β = 0.165	r = -0.072 p = 0.635 1-β = 0.077

4.3.3 2D:4D and lateralization for counting vs. subitizing – ANOVA analysis

As described above, four way ANOVAs were used to explore any overall main effects of digit ratio (low vs. high; separate ANOVAs conducted for groups formed on the basis of either right or left hand measures), sex (male vs. female), process (subitizing vs. counting), and visual field (right vs. left) and their possible interactions on RTs. In light of the revealed violations of normality however the results of this analysis should be viewed with caution.

Given the large number of main and interaction effects possible from 4-way ANOVA analysis the most pertinent results in the context of the current thesis, i.e. those relating to a potential effect of 2D:4D (both main and interaction effects) will be reported first. Any further significant main effects will then be reported followed by any significant interaction effects not involving the factor of 2D:4D.

4.3.3.1 Analysis including digit ratio groups split on the basis of right hand 2D:4D measures

Analysis including right hand 2D:4D data showed no significant main effect of digit ratio and only one significant interaction effect involving the factor of digit ratio, namely a three-way interaction effect between the factors digit ratio group, process (subitizing vs. counting) and visual field, see table 10.

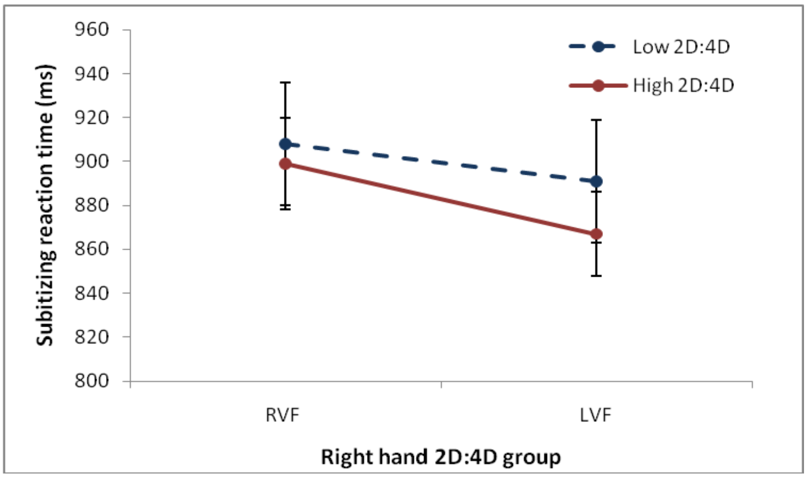
Table 10

F, p, df, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of 2D:4D (significant effect highlighted in bold).

Effect	**F value**	**df**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.045	1, 72	0.832	203585.7	0.001	0.055
2D:4D x Process interaction	0.012	1, 72	0.913	55663.51	0.0002	0.051
2D:4D x Visual field interaction	0.365	1, 72	0.548	2783.699	0.005	0.092
2D:4D x Sex interaction	0.08	1, 72	0.778	203585.7	0.001	0.059
2D:4D x Process x Sex interaction	0.028	1, 72	0.867	55663.51	0.0004	0.053
2D:4D x Visual field x Sex interaction	0.001	1, 72	0.975	2783.699	0.00001	0.05
2D:4D x Process x Visual field interaction	**4.062**	**1,72**	**0.048**	**3224.823**	**0.053**	**0.511**
2D:4D x Process x Visual field x Sex x interaction	2.773	1, 72	0.1	3224.823	0.037	0.376

As can be seen in figure 4a, the effect of visual field on subitizing appears to be similar for both high and low 2D:4D participant such that both groups displayed a left visual field advantage for the task. This effect of visual field appears to be slightly more pronounced for high 2D:4D participants. As can been in figure 4b, however a similar pattern of results was not observed for counting RTs where low 2D:4D (high PT) participants demonstrated a slight left visual field advantage and high 2D:4D participants demonstrate a slight right hemisphere advantage. In order to explore the interaction between process (subitizing vs. counting), visual field, and digit ratio group further the analysis was broken down by process so to examine individual or combined effects of visual field and right hand digit ratio group separately for the two enumeration strategies of counting and subitizing. Post-hoc analysis revealed a significant LVF advantage for subitizing RT $F_{(1,74)} = 21.094$, MSe =1042.321, $p < 0.001$, $\eta p^2 = 0.222$, $1-\beta = 0.995$, see table 7. A significant effect of visual field however was not displayed for counting RT, $F_{(1,74)} = 0.061$, MSe = 4970.913, $p = 0.805$, $\eta p^2 = 0.001$, $1-\beta = 0.057$. No significant main effects of 2D:4D or 2D:4D x visual field interaction effects were revealed for either subitizing (main effect - $F_{(1,74)} = 0.234$, MSe = 42828.114, $p = 0.63$, $\eta p^2 = 0.003$, $1-\beta = 0.076$, interaction – $F_{(1,74)} = 2.084$, MSe = 1042.321, $p = 0.153$, $\eta p^2 = 0.027$, $1-\beta = 0.297$) or counting (main effect – $F_{(1,74)} = 0.046$, MSe = 216857.514, $p = 0.83$, $\eta p^2 = 0.001$, $1-\beta = 0.055$, interaction – $F_{(1,74)} = 1.724$, MSe = 4970.913, $p = 0.193$, $\eta p^2 = 0.023$, $1-\beta = 0.254$) RTs.

a)



b)

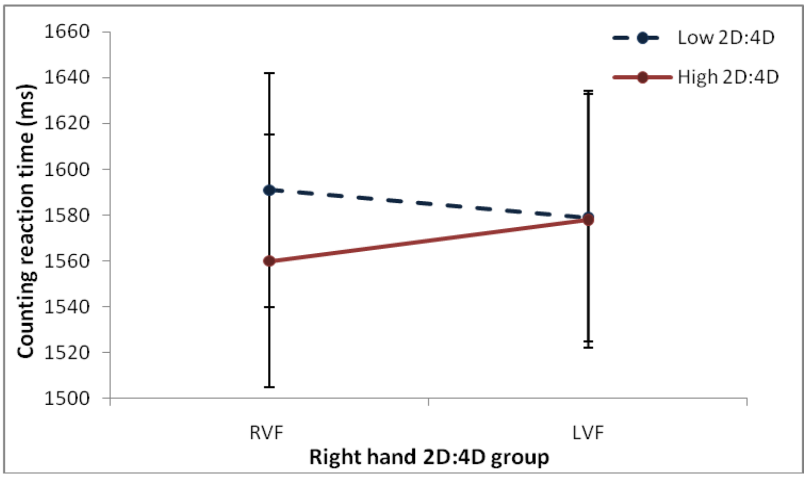


Figure 4 a & b. Mean subitizing (a) and counting (b) RT in both low and high 2D:4D participants for information presented to both the left and right visual fields.

No significant main effects of sex or visual field were indentified in the overall 4-way ANOVA. 4-way ANOVA analysis however revealed a significant overall main effect of process (subitizing vs. counting) on RTs, $F_{(1,72)} = 586.659$, MSe = 55663.51, $p < 0.001$, $\eta p^2 = 0.891$, $1\text{-}\beta = 1$. As can be seen in table 7, responses to quantities in the subitizing range were significantly faster and more accurate than those to quantities in the counting range.

4-way ANOVA analysis also revealed a significant two-way interaction between visual field and process, $F_{(1,72)} = 4.56$, MSe = 3234.823, $p = 0.036$, $\eta p^2 = 0.06$, $1\text{-}\beta = 0.558$, with a left visual field (LVF; right hemisphere) advantage revealed for subitizing and a right visual field (RVF; left hemisphere) advantage revealed for counting. As earlier post-hoc analysis of the three way interaction between 2D:4D group, process and visual field already revealed a significant left visual field advantage for subitizing in the absence of any significant effect of visual field on counting RT, and given that any main effect of process on RT to the separate visual fields was not of specific interest to the current experiment, no further post-hoc analysis was conducted to explore this two way interaction.

Finally, a significant interaction was revealed between the factors, enumeration process and sex, $F_{(1,72)} = 5.494$, MSe = 55663.509, $p = 0.022$, $\eta p^2 = 0.071$, $1\text{-}\beta = 0.638$. While males and females demonstrated similar subitizing RTs with a very slight female advantage (males mean = 896.5ms, SD = 197.28; mean females = 888.39ms, SD = 101.08), males showed faster RTs than females to numerosities in the counting range (males mean = 1503.58ms, SD = 371.71; mean females = 1624.77ms, SD = 288.86).

Post-hoc t-tests (Bonferroni corrected for multiple comparisons, $\alpha = 0.0125$) however revealed no significant sex differences for either subitizing or counting RTs. Similar to results for analysis of the whole sample, both males, $t_{(29)} = 14.26$; $p < 0.001$, and females, $t_{(45)} = 21.48$, $p < 0.001$, considered separately demonstrated significantly faster RTs to quantities in the subitizing range as compared to the counting range.

4.4 Discussion

The objective of the present experiment was to explore any potential relationships between 2D:4D as a proxy marker of PT, the basic enumeration processes of subitizing and counting, and cerebral lateralization for these two processes. A number of important limitations however were identified in the methodology which questions the extent to which the process of subitizing may have been evaluated. The results section of the current chapter therefore reports ANOVA findings only relating to the analysis of right hand 2D:4D data. As the methodological limitations limit the extent to which clear conclusions can be reached the additional reporting of left hand 2D:4D ANOVA results can add little to the interpretation of the findings. F, p, MSe, effect size (partial eta squared - ηp^2), and power (1-β) values relating to 4-way ANOVA analysis including the factors process (subitizing vs. counting), visual field, sex and 2D:4D group formed on the basis of left hand 2D:4D data however can be viewed in Appendix 2.

In line with evidence for established sex differences in 2D:4D, the present study revealed significant sex differences in right hand 2D:4D measures. As anticipated, males demonstrated significantly lower 2D:4D ratios, assumed to reflect higher PT exposure, than females (Lutchmaya et al., 2004; Manning et al., 1998; Manning, 2002). Notably however, although in a similar direction, sex differences in left hand 2D:4D measures did not reach significance. This finding may not be particularly unusual in light of evidence that while sex differences in 2D:4D are present in both hands, the sexual dimorphism may be larger on the right hand as compared to the left (Manning et al., 1998; Williams et al., 2000). The precise reason for this remains unknown.

Data from the current study demonstrated a significant process (subitizing vs. counting) x sex interaction. While males and females achieved similar RTs on the subitizing task, male RTs on the counting task were faster than those demonstrated for females. Post-hoc analysis however revealed no significant sex differences between either subitizing or counting performance analysed independently. This finding suggest that patterns of performance between the sexes may differ depending upon which task is being evaluated, the results however are in accordance with previous evidence in children which suggests an absence of sex differences for basic mathematical skills (Geary, 1996; Fink et al., 2006).

Results from the present experiment also found a significant process x visual field interaction, whereby a LVF (right hemisphere) advantage was revealed for the process of subitizing, and a RVF (left hemisphere) advantage was observed for the process of counting. Such findings are consistent with those reported by Pesini and Tessari (2001) who demonstrated similar differences in hemisphere specialization for the two processes as the current experiment. Post-hoc analysis however revealed that visual field differences were only significant for subitizing. This is in contrast to Pasini and Tessari (2001) who reported stronger visual field advantages for the process of counting. The revealed LVF (right hemisphere) advantage in the current study is also similar to alternative behavioural evidence (Pasini & Tessari, 2001; Arp et al., 2006; Jackson & Coney, 2004; Boles et al., 2007) but contrary to Butterworth (1999) who suggested that subitizing may actually be sub served by the left hemisphere and brain imaging evidence which typically reports no particular hemispheric bias for the process of subitizing (Sathian et al., 1999; Piazza et al., 2002; Nan et al., 2006).

Unfortunately however these results should be viewed with caution as the present study possessed a clear confound with regard to method of manual response which is likely to have biased interactions between visual field and process (subitizing vs. counting). As numbers on the response box were presented in a linear fashion, all quantities in the subitizing range were responded to using the right hand, with all quantities in the counting range responded to with the left hand. The significant right visual field advantage for subitizing therefore may be explained with reference to this response pattern (faster responses being a result of a correspondence between visual field presentation and response hand). Nevertheless, there is no *a priori* reason to expect the effects of this confound to selectively influence quantities in the subitizing range, given the lack of significant findings with regard to counting it is unlikely that the method of manual response can entirely account for the identified pattern of results.

Correlation analyses found no significant relationships between either right or left hand 2D:4D and subitizing or counting performance, overall or in male and female data analysed independently. Similarly ANOVA analysis revealed no main effects of either right or left hand 2D:4D or process (subitizing vs. counting) x 2D:4D interactions. The result of the current experiment therefore are contrary to evidence demonstrating associations between 2D:4D and aspects of numerical and mathematical performance in children and adults (Bull & Benson, 2006; Fink et al., 2006; Kempel et al., 2005; Luxen & Buunk, 2005). However, while the lack of significant effects with reference to subitizing initially appears contrary to the suggestion that PT may influence our apparently innate 'core' numerical system for the precise representation of small quantities, it is vital to highlight that typical subitizing response patterns were not evident in the current data. As mentioned previously, results from the present study displayed elevated subitizing response times and a steeper subitizing RT x set size function than those typically reported in prior literature (e.g. Trick & Pylyshyn, 1994; Dehaene & Cohen, 1994; Mandler & Shebo, 1982).

Critically, in an attempt to tap both processes of counting and subitizing the present study offered unlimited response time during the enumeration tasks. As participants were given the opportunity to adopt either strategy of enumeration therefore it is possible that the uncharacteristic data patterns revealed in the current study towards dot patterns in the range of 2-4 could potential imply the use of a counting as opposed to subitizing in a number of participants. Such an effect has been documented in 3-year-old children where a spontaneous preference for the use of a counting strategy over subitizing was observed for set sizes 2-4 under conditions that did not restrict response time (Silverman & Rose, 1980). As it is possible then that the process of subitizing was not effectively assessed in all participants, a potential association between 2D:4D and subitizing cannot be entirely dismissed on the basis of present data.

It is also important to highlight that revealed statistical power was low across all analyses (for both correlation and ANOVA calculations). Based on the conventions described by Cohen (1988) previous correlational data from Fink et al. (2006) generally shows medium to high effect sizes for a relationship between 2D:4D and numerical performance in males and small to moderate effect sizes for a relationship between 2D:4D and numerical performance in females. Similarly correlational data reported by Brosnan (2008) reports moderate effect sizes for a relationship between 2D:4D and numeracy in males and small effect sizes for a relationship between 2D:4D and numeracy in females. Utilising similar sample sizes to those employed in the present study both Fink et al. (2006) and Brosnan (2008) report significant associations between 2D:4D and numerical ability in males but not females. In the present study the majority of effect sizes relating to the ANOVA analysis were very low. Similarly correlation values for analysis of all data were below the classification for a small effect size.

Correlation analysis of male and female data separately however suggested the presence of a number of small-to-moderate effects. Where such effects were observed the direction of the relationships suggested positive associations between left hand 2D:4D and subitizing and counting reaction times in males and negative associations between right and left hand 2D:4D and subitizing reaction times in females ad well as negative associations between right hand 2D:4D and counting reaction times in females. Such effects are in line with previous research which suggests sex dependent relationships between potential measures of PT and mathematical performance (Brosnan, 2008; Fink et al., 2006; Finegan et al., 1992; Kempel et al., 2005), and the findings of Fink et al. (2006) and Brosnan (2008) who report lower 2D:4D to be associated with improved performance in males.

Prospective power analysis using G*Power 3.1.2 (Faul et al., 2009) suggest that 193 participants are required in order to achieve a power of 0.8 using correlation for a small to moderate effect size (entered as 0.2). When using Spearman's this required sample size rises to 213 according to the calculations described in chapter 2 (section 2.3.1, p51) For a 2x2x2x2 ANOVA with 2 independent measures variables and 2 repeated measures variables in order to achieve a power of 0.8 for a small to moderate effect size (entered as d = 0.175) 260 participants are required (see Appendix 16 for the G*Power output relating to these calculations). This calculation however is based on the assumption that the design is balanced, unbalanced designs such as those used in the current study may reduce power. It is important to recognise therefore that the lack of significant findings relating to 2D:4D in the current experiment may be consequence of low power.

While no direct 2D:4D effects were revealed for performance on the experimental task, a significant right hand 2D:4D x visual field x process interaction was revealed which implied different patterns of lateralization for high vs. low 2D:4D participants between the process of subitizing vs. counting. While further analysis did not reveal any further effects of 2D:4D on the processes of subitizing and counting analysed separately the interaction cannot easily be accounted for with regard to the methodological limitations and thus requires further investigation in order to explore and clarify any potential relationship between 2D:4D and lateralization for these two basic enumeration process.

Although the finding regarding a potential association between 2D:4D and lateralization for the processes of subitizing and counting is intriguing and requires further investigation, given the identified methodological limitations of the present study, speculation as to the implications of the results for the potential effect of PT on patterns of lateralization on the basis of current data would be premature for a number of reasons. Firstly, it is important to highlight that normality was violated for measures of subitizing reaction times, this is likely to have effected the efficiency of the ANOVA analysis. As noted in the results sectiontherefore, any effect identified via ANOVA analysis should be viewed with caution.

Secondly, it is impossible to rule out potential group differences in the adopted enumeration strategy for quantities in the subitizing range. Such potential differences may have presented important effects on revealed findings that cannot be controlled. Similarly, it is unclear whether the same findings would have been obtained had method of manual response to quantities in the subitizing and counting ranges been adequately controlled. Again, as all numerosities in the range of 2-4 were responded to using the left hand, and all numerosities in the range of 6-8 responded to using the right hand, one cannot dismiss the possibility that individual and group difference in susceptibility to stimuli-response compatibility effects (where the time taken to respond to a stimuli is associated with the spatial relation between stimuli position and side of response) could have heavily influence present findings.

Beside such methodological problems one further limitation with reference to the present study is the fact that generic lateralization to speeded response was not adequately controlled. Firstly, the tasks included to measure general RT (both non-numerical and numerical) were designed to be employed as covariates in the analysis, upon considering the required 4-way ANOVA however it became apparent that, as general RT is fundamentally related to the source of variance between the processes of subitizing vs. counting and potentially between RT to the left vs. right visual field, the use of covariates controlling for general RT when considering such factors would be inappropriate (see Miller & Chapman, 2001). Such general RT measures therefore could only be justifiably used as covariates when more closely considering any revealed pre-existing group differences (sex or 2D:4D group differences) in RT during the process of either subitizing or counting. As, given the results, a closer investigation of such group differences was never warranted the controls for general RT were never actually utilised. Secondly, both tasks adopted to assess generic RT during a speeded response task utilised central stimuli presentation. Thus, while any different patterns of lateralization according to 2D:4D group that might have been revealed may have indicated a possible effect of *in utero* testosterone concentration on subitizing and/or counting it is equally possible that similar findings could have been obtained utilising similar, non-numerical RT tasks. This would suggest a more general effect of PT on lateralization for different speeded response tasks, as opposed to a more specific relationship between *in utero* testosterone exposure and lateralization for the processes of subitizing and counting. At present on the basis of the current experimental design it would have been impossible to tease apart the level at which testosterone may be effecting lateralisation on the task.

Future studies attempting to re-explore associations between 2D:4D or indeed any potential measure of PT and lateralization for the basic enumeration strategies of subitizing and counting should attempt to usefully address the methodological limitations identified in the current study. Perhaps most importantly, a method of manual response should be properly controlled in order to remove any potential confound which may exist as a result of response hand bias. Secondly, in order to be confident that a subitizing strategy has been induced it appears necessary to limit presentation and response times making the use of a counting strategy impossible. Finally, controls for generic RT should incorporate an identical procedure with regard to stimulus presentation as the experimental task, e.g. where lateralization is being examined, both experimental and control tasks should similarly incorporate lateralized stimulus presentation so that more general relationships with lateralization for any similar speeded response task can be taken into account. In an attempt to address the issue regarding the use of a general RT covariate that is intimately associated with the source of variance between one of the factors to be included in the analysis it may be preferable to include general RT as an additional independent variable so that any interaction between the effects of 2D:4D on the experimental task and 2D:4D on a control task can be explored. An interaction would imply a different effect of 2D:4D for the experimental task vs. a control task and thus would allow for a consideration of the possibility that any associations between the basic numerical task and 2D:4D could potentially be secondary to a more general relationship between PT and performance.

In summary, the current experiment revealed no direct relationships between 2D:4D and performance on an enumeration task assessing response times to quantities in both the subitizing and counting ranges. Findings from the current study however do suggest evidence for possible novel and interesting relationships between 2D:4D and lateralization for subitizing and counting. Unfortunately the study possessed a number of important methodological confounds which make any confident conclusions or speculation regarding the results difficult. Further research addressing these methodological limitations is necessary in order to re-examine relationships between 2D:4D and the basic enumeration strategies of counting and subitizing.

Chapter 5

Experiment 3: Re-Examining the Association between 2D:4D and Subitizing in Adults.

5.1 Introduction

Utilising 2D:4D as a somatic marker of prenatal testosterone (PT), experiment 2 sought to; *a)* explore potential correlations between PT and one aspect of 'core' numerical processing (i.e. subitizing), *b)* compare and contrast this potential relationship with possible associations between 2D:4D and a similar basic numerical skill but one that is not typically identified to be innate (i.e. counting), and *c)* based on hypotheses that PT may modify cerebral lateralization (see chapter 1), compare and contrast potential relationships between 2D:4D and two numerical processes which may potentially to show different patterns of hemispheric specialization (see chapter 4). Whilst findings highlighted some potential associations between 2D:4D and lateralization for the two basic process of enumeration, no evidence for a direct relationship between 2D:4D and either subitizing or counting was revealed. The study however did possess a number of key methodological limitations, making any clear conclusions and interpretations of the findings difficult. In an attempt to re-examine possible relationships, taking prior methodological limitations into account, the current study presents a modified partial replication of experiment 2.

In attempting to assess both processes using one task with similar response characteristics for quantities in both the subitizing and counting ranges, it appears that, during experiment 2 a process of counting as opposed to subitizing may have been utilised for the enumeration of dot patterns 2-4. As one of the primary aims of the study was to explore relationships between 2D:4D and 'core' numerical processing, the current study will simplify the procedure to focus specifically on subitizing, thus the enumeration of quantities in the counting range will not be assessed. This will facilitate the allocation of a suitable response time to ensure counting is not employed as an alternative strategy for the enumeration for quantities 2-4.

As counting won't be assessed, and thus possible differences in 2D:4D relationships depending upon hemispheric specialization for the two processes of counting and subitizing cannot be evaluated, the procedure of the current study will be further simplified (in relation to experiment 2) by adopting centrally presented stimuli as opposed to lateralized stimuli presentation. Given the preliminary nature of this area of research, this method was chosen to allow for a specific focus on relationships between 2D:4D and general performance for subitizing, without the additional factor of lateralization in the task. This will ultimately offer a more direct assessment of any relationship which may exist.

The previous study failed to adequately control for the method of manual response, resulting in all quantities in the subitizing range being responded to using the left hand and all quantities in the counting range being responded to using the right hand. This method of response may have exerted important biases in patterns of hemispheric specialization for the experimental task. In the present experiment, in order to be confident that individual differences in response patterns are unlikely to be influenced by hand of response, response hand will be counterbalanced across participants.

A final limitation of the previous study was its failure to adequately control for general reaction time and thus the possibility that any potential associations between 2D:4D and performance on the experimental task are independent of basic speed of processing. As the presentation of stimuli during control tasks was not lateralized, it would have been impossible in the previous study to have established whether revealed relationships between 2D:4D and subitizing or counting were specific to these two processes or simply a result of a more general relationship between 2D:4D and lateralized patterns of response during reaction time tasks. As presentation of experimental stimuli in the current experiment will not vary according to visual field, a generic assessment of lateralization for speeded response is not required. An appropriate comparable control task however employing an identical set-up to the experimental task with regard to stimuli presentation and method of response will be included. In addition, in experiment 2 the factor of general reaction time was designed to be included as a covariate in the analysis in order to control for basic speed of processing. This method however was inappropriate given that the factor was fundamentally related to the source of variance between two of the independent variables. In order to avoid this issue in the current study, rather than employing general reaction time as a covariate the potential role of general reaction time was considered by exploring any interaction that might exist between the effect of 2D:4D and the factor of task (Control vs. Experimental). This allows for the consideration of potential difference in the effects of

2D:4D between performance on a general reaction time task in contrast to a core numerical task.

In the previous experiment both a non-numerical and numerical control task were utilised. In the current study however, in deciding to use a control task with an identical procedure for stimulus presentation and method of manual response, it is possible that any numerical control task may potentially tap into a subitizing representation. There is evidence, for example, that during numerical comparison Arabic digits are converted from their symbolic format to a quantity representation (Dehaene, 1996). It is possible therefore that a subitizing effect may be accessed during all forms of numerical recognition, regardless of input notation. It was decided therefore that a non-numerical control would be most suitable to the present experiment.

In contrast to the previous experiment the present study also employed an alternative method of 2D:4D measurement. In both experiment 1 and 2 measurements of the second and fourth fingers were taken directly from the hands, in the current study however scans of the hand were employed in order to calculate 2D:4D. The procedure was amended for two primary methodological reasons, 1) the adoption of a hand scans is less time consuming than directly measuring fingers during the experimental procedure and thus reduces testing time and 2) the use of hand scans provides a permanent facsimile of the hand which can be accurately measured and, if necessary, re-measured subsequent to testing. While there is some evidence that the use of photocopies may yield lower ratios and may show a stronger sex difference than direct finger measurements (Manning et al., 2005) intraclass correlation coefficients between measures of 2D:4D derived from the two different techniques are reported to be high.

Similar to the predications for experiment 2, based on evidence from Bull and Benson (2006) that 2D:4D may relate to one aspect of numerical processing thought to partially reflect core numerical skill, it is hypothesised that lower (more masculinised) 2D:4D ratios will relate to increased automaticity for the primitive process of subitizing and, thus, decreased RTs during a subitizing task. It is expected that this relationship will not exist purely as a consequence of a more general relationship between 2D:4D and general reaction time, it is thus anticipated that any association between 2D:4D and performance on the subitizing task will be greater than, or in contrast to, any relationship between 2D:4D and performance on the control task. Again, given inconsistencies in previous literature regarding the presence of significant effects depending upon sex (see chapter 1), it is predicted that sex differences may exist in the strength and, perhaps, the nature of associations. Similar to experiment 2 therefore it is the specific interactions between 2D:4D, task (subitizing vs. control) and sex which are of particular interest. In order to evaluate any possible interactions therefore participants were again categorised into high vs. low 2D:4D groups on the basis of their raw 2D:4D measurements so to facilitate ANOVA analysis.

5.2 Method

5.2.1 Design

The experiment adopted a 2 x 2 x 2 x2 mixed measures, quasi-experimental design including the four factors of; 2D:4D (low vs. high – determined according to median split), sex (males vs. females), task (subitizing vs. control) and response hand (right vs. left). Basic associations between right and left hand 2D:4D and subitizing performance were also explored using correlations.

5.2.2 Participants

Eighty self-reported heterosexual participants gave their written informed consent to partake in the experiment. A total of 40 males and 40 females were recruited from the student population of Northumbria University, and paid £3 for participation. All participants were right handed according to self reported writing hand (assessed by means of a single biographical question). All participants reported a heterosexual sexual orientation, normal or corrected to normal visual acuity, no present or previous injury to the second or fourth fingers of either hands, and no known hormonal abnormalities or the taking of any hormone influencing drugs (excluding the contraceptive pill). Adopting the exclusion criteria of experiment 2, participants not of the majority ethnicity (white British) for the sample were removed from the analysis (n = 10) resulting in a final sample of 34 males (mean age = 23.62 years, SD = 5.58) and 36 females (mean age = 22.19 years, SD = 5.7).

5.2.3 Measures

5.2.3.1 Reaction time tasks

The method of creation and presentation (including the adopted PC and monitor, dot size, visual parameters of dot pattern presentation and minimum and maximum distance between dots) of the experimental stimuli in the current experiment was identical to that adopted in the previous experiment (see figure 2a), with the exception that the stimuli in the present study consisted only of the quantities 2, 3 and 4 presented to the centre of the screen. Eight sets of different dot patterns were randomly generated for each quantity, and presented in a pseudo-random order such that each quantity appeared a total of 25 times.

Prior to every trial, a central fixation point (X) appeared on the screen for a duration of 1000ms. Stimuli were presented 1000ms after fixation for a total of 50ms, followed by a blank screen. Participants were given 1500ms in which to identify the number of dots presented. Responses were measured to the nearest millisecond via a three-button response pad with keys labelled '2', '3', and '4'. Response deadline was signalled with a single beep, after which, the inter-trial-interval was randomly set between 1000ms and 1500ms. Each participant began with a practice block of 20 trials.

As a control for general reaction time, all participants performed a simple colour recognition task in which a 7.5cm × 7.5cm coloured square (either red, yellow, or blue) was presented to the centre of the screen using similar parameters of duration and presentation as the experimental task. Participants were instructed to identify the colour presented using coloured keys on the response box. For this control task each participant completed a practice block of 10 trials followed by a main block of 30 trials (each colour was presented 10 times in a random order).

The order in which the tasks were completed was counterbalanced. Participants completed the whole programme once using the index finger of their left hand, and once using the index finger of their right hand. Order of hand used was also counterbalanced such that half of the participants completed the programme using their left hands first while the remainder completed the programme using their right hands first. An equal number of males and females were allocated to the two orders of response hand groups.

5.2.3.2 Second to fourth finger ratio measures

The lengths of the 2nd and 4th digits were measured from the basal crease to the tip on the ventral surface of both left and right hands using colour flatbed-scanned images (resolution 300 dpi). In order to ensure measurement repeatability, two separate images of each hand were obtained, once prior to conducting the experiment and once following completion. An average length of 2nd and 4th fingers was calculated and 2D:4D computed using these averaged measurements. Participants were instructed to place the palm of their hand in a relaxed position, with fingers evenly spaced on the glass bed of the scanner without applying pressure. Hand scans were zoomed to 200% original size and printed images were measured using Vernier Callipers accurate to 0.01 mm. Intraclass correlation coefficients (r_I) showed high retest-reliability between first and second measurements of both the right (second r_I = 0.983; fourth r_I = 0.988; 2D:4D r_I = 0.947) and left hands (second r_I = 0.992; fourth r_I = 0.991; 2D:4D r_I = 0.932). From initial and final second and fourth finger measurements the technical error of measurement (TEM) and relative technical error of measurement (rTEM) were computed according to protocol established by Weinberg et al., (2005). For second digit measurements TEM was computed to be 0.719 for right hand measures and 0.477 for left hand measures with the corresponding rTEM calculated to be 0.977% and 0.656% respectively. TEM for fourth digit measurements was calculated at 0.619 and 0.527 and rTEM at 0.819% and 0.698% respectively. As stated in chapter 2, Weinberg et al. (2005) recommends a cut off point of 5%, with all rTEM percentages above this considered imprecise. Again, in accordance with this published criterion, an acceptable degree of precision for second and fourth finger measurements was achieved.

As anticipated, males demonstrated lower ratios (right hand mean = 0.966, SD = 0.03; left hand mean = 0.958, SD = 0.0331), assumed to reflect higher PT, as compared to females (right hand mean = 0.98, SD = 0.026; left hand mean = 0.97, SD = 0.026), sex differences in 2D:4D however were only significant for the right hand, $t_{(68)}$=2.182; p=0.033 (left hand – $t_{(68)}$=1.793; p = 0.077).

In order to facilitate investigation of the possible interaction between factors included in the experiment, within-sex median splits according to 2D:4D for both the right (median males = 0.964, median females = 0.978) and left (median males = 0.957, median females = 0.975) hands were applied to the data, resulting in the formation of two sets of high vs. low 2D:4D groups. Mean 2D:4D values were confirmed to be significantly different between these groups, see appendix 3.

5.2.4 Procedure

The study was approved by the School of Psychology & Sport Sciences Ethics Committee. Following informed written consent, participants completed a brief biographical questionnaire, and then initial hand scans were obtained. For the computerised tasks, participants were requested to place their head on a chin rest positioned approximately 50cm from the centre of the monitor and focus their gaze towards the centre of the screen. Participants were instructed to respond as fast as they could while also paying attention to accuracy. Participants completed the experimental and control tasks, and reaction times and errors were recorded. Testing on the computer took approximately 30min after which final hands scans were taken. On completion, participants were fully debriefed.

5.3 Results

5.3.1 Behavioural data

Overall means and standard deviations for reaction times (RT) to correct responses and percentage error for subitizing and control tasks are presented in table 11. Figure 5 illustrates separate RTs and percentage error scores for responses to the quantities 2, 3, 4 and the colours red, yellow and blue. Full tables of means for right and left hand 2D:4D data and subitizing and control task performance, overall and in both males and females considered separately, can be viewed in Appendix 3. Assessment of normality according to one-sample Kolmogorov-Smirnov analysis revealed that all reaction time measures on both the subitizing and control task were normally distributed. Error scores on both tasks however were revealed not to be normally distributed, see Appendix 3.

Table 11

Means and standard deviations for participant (n = 70) reaction times in ms to correct responses and percentage error for subitizing and colour recognition tasks.

	Percentage Error - Mean(SD)			***RT (ms) - Mean(SD)***		
Task	*LH Response*	*RH Response*	*Overall*	*LH Response*	*RH Response*	*Overall*
Subitizing	4.02(3.27)	4.11(3.59)	4.07(2.98)	761(74.09)	730(74.87)	746(71.33)
Colours	2.52(3.38)	2.76(3.54)	2.64(2.62)	551(72.46)	533(79.2)	541(69.83)

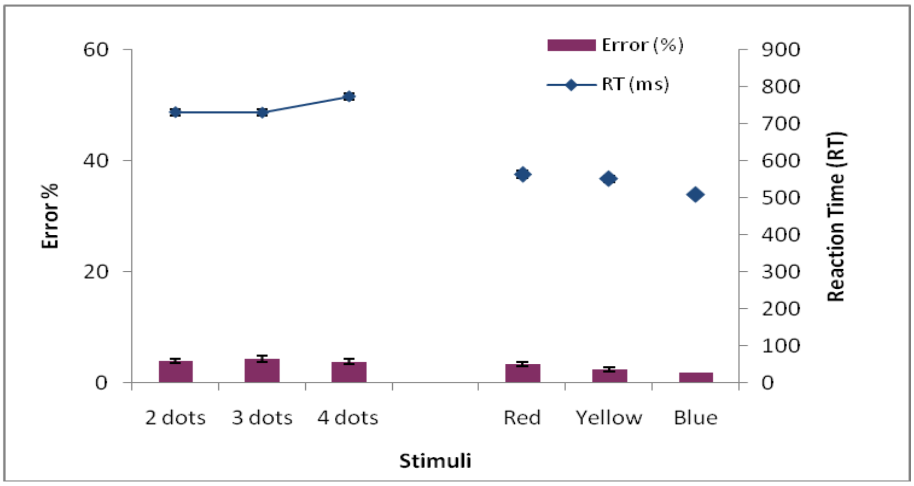


Figure 5. Mean reaction times and percentage error scores for the enumerations of quantities (2-4) in the subitizing task and colours (red, yellow and blue) in the colour recognition task, error bars indicate SEM.

Paired samples t-tests (Bonferroni corrected for multiple comparisons = 0.025) were conducted in order to explore differences in reaction time as a function of increasing numerosity. As percentage error variables were not normally distributed the same analysis (i.e. exploring changes as a function of increasing numerosity) for percentage error scores were conducted using the Wilcoxn signed ranks test. No significant response time differences were observed between identification of the quantities 2 and 3. Average reaction times to the enumeration of 3 dots was actually marginally lower than average reaction time to the enumeration of 2 dots (mean difference = 1.24ms). Paired sample t-tests however revealed reaction times to 3 dots to be significantly faster than averaged 4 dot response times, $t_{(69)} = -6.764$ $p<0.001$ (mean difference = 43.9ms). No significant differences in percentage error were apparent between the identification of 2 and 3 (mean difference = 0.43) and 3 and 4 dots.

Average 4 dot percentage errors scores were also lower than average 3 dot percentage error scores (mean difference = 0.49). With regard to the control task, paired sample t-tests (Bonferroni corrected for multiple comparisons = 0.017) revealed significant response time differences between the colours; red and blue; $t_{(69)} = 7.261$, $p<0.001$, and yellow and blue; $t_{(69)} = 5.605$, $p<0.001$. As can be seen in figure 5, the fastest reaction times were observed for the colour blue, followed by yellow with the slowest average responses observed for identification of the colour red. Significantly higher average percentage error was also revealed for the identification of the colour red as compared to blue, $Z = -2.438$, $p = 0.015$.

No significant speed/accuracy associations were revealed for performance on either the colour recognition or subitizing tasks. As can be seen in table 11, percentage errors for both the control and subitizing task did not exceed 5%, similar to the previous experiment therefore analysis in the current experiment focused exclusively on reaction time data.

5.3.2 Correlations

Pearson's correlations were conducted in order to explore simple associations between 2D:4D and performance on the subitizing task, see table 12. Power calculations (computed using G*Power, Faul et al., 2009) for each analysis were again very low. The majority of effect sizes were also quite low although a number of coefficient values did suggest small to moderate effects. Small to moderate effects for example were observed between right hand 2D:4D and left hand subitizing performance in males (positive direction, i.e. low 2D:4D associated with faster reaction times), and right and left hand 2D:4D and subitizing overall and left hand subitizing performance in females (negative direction, i.e. low 2D:4D associated with slower reaction times). Again however, no significant correlations were revealed for subitizing responses in the overall sample or male and female data analysed independently.

Table 12

Pearson's correlation coefficients (r) for the relationship between left and right hand 2D:4D and subitizing performance. P values and power calculations (1- β) for each analysis are also listed.

	Males and Females n = 70		**Males n = 34**		**Females n = 36**	
	Right 2D:4D	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Subitizing Overall	$r = -0.009$ $p = 0.943$ $1\text{-}\beta = 0.051$	$r = -0.094$ $p = 0.439$ $1\text{-}\beta = 0.121$	$r = 0.056$ $p = 0.751$ $1\text{-}\beta = 0.061$	$r = -0.063$ $p = 0.725$ $1\text{-}\beta = 0.064$	$r = -0.108$ $p = 0.53$ $1\text{-}\beta = 0.96$	$r = -0.15$ $p = 0.383$ $1\text{-}\beta = 0.141$
Subitizing Left Hand	$r = 0.006$ $p = 0.96$ $1\text{-}\beta = 0.05$	$r = -0.083$ $p = 0.496$ $1\text{-}\beta = 0.105$	$r = 0.136$ $p = 0.442$ $1\text{-}\beta = 0.119$	$r = -0.035$ $p = 0.845$ $1\text{-}\beta = 0.054$	$r = -0.209$ $p = 0.221$ $1\text{-}\beta = 0.233$	$r = -0.185$ $p = 0.281$ $1\text{-}\beta = 0.191$
Subitizing Right Hand	$r = -0.022$ $p = 0.853$ $1\text{-}\beta = 0.054$	$r = -0.097$ $p = 0.423$ $1\text{-}\beta = 0.126$	$r = -0.023$ $p = 0.898$ $1\text{-}\beta = 0.052$	$r = -0.086$ $p = 0.63$ $1\text{-}\beta = 0.077$	$r = 0.014$ $p = 0.938$ $1\text{-}\beta = 0.051$	$r = -0.093$ $p = 0.589$ $1\text{-}\beta = 0.084$

5.3.3 2D:4D, sex and subitizing

A four-way ANOVA was used to explore any overall main effects of digit ratio (low vs. high), sex (male vs. female), task (subitizing vs. control), and response hand (left vs. right) and their possible interactions on RTs. Results from the current experiment will be presented in accordance with the previous chapter, i.e. with those relating to a potential effect of 2D:4D reported first.

5.3.2.1 Analysis including right hand 2D:4D

No significant main effect of 2D:4D on reaction times was identified. A significant 4-way interaction however was revealed between all factors, right-hand 2D:4D, task, sex, and response hand, see table 13. No further significant interactions involving the factor of right hand 2D:4D were found to be significant. Results of the four-way ANOVA revealed a significant overall main effect of task on reaction times $F_{(1, 66)} = 1422.29$, MSe = 2040.06, $p < 0.001$, $\eta p^2 = 0.956$, $1\text{-}\beta = 1$. Responses to the colour recognition task were significantly faster and more accurate than those observed for subitizing (see table 11). A significant overall main effect of hand on reaction times was also revealed, $F_{(1,\ 66)} = 28.443$, MSe = 1424.047, $p < 0.001$, $\eta p^2 = 0.301$, $1\text{-}\beta = 1$. Given that all participants were self-reported right handed, unsurprisingly, the result showed a significant overall right hand advantage (see table 11).

Table 13

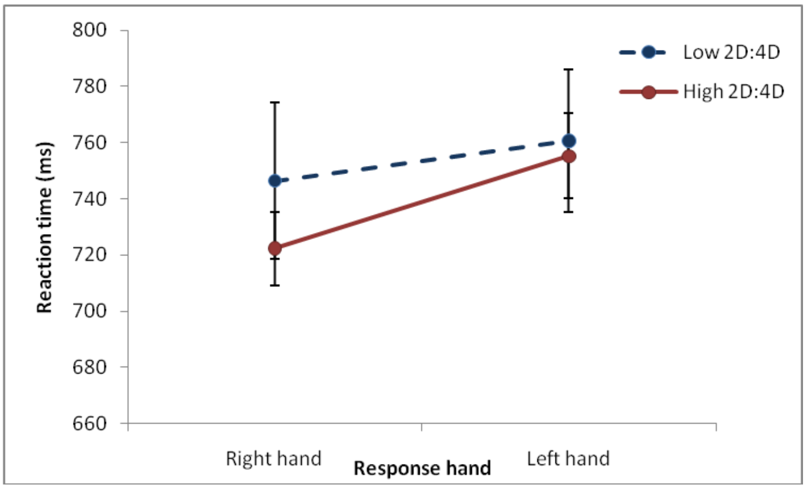
F, p, df, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of right hand 2D:4D (significant effect highlighted in bold).

Effect	**F value**	**df**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.716	1,66	0.4	18461.25	0.011	0.133
2D:4D x Task interaction	1.251	1,66	0.267	2040.06	0.019	0.197
2D:4D x Sex interaction	0.001	1,66	0.975	18461.25	0.00002	0.05
2D:4D x Response hand interaction	2.25	1,66	0.138	1424.047	0.033	0.315
2D:4D x Task x Sex interaction	1.478	1,66	0.228	2040.06	0.022	0.224
2D:4D x Response hand x Sex interaction	0.03	1,66	0.863	1424.047	0.0005	0.053
2D:4D x Task x Response hand interaction	0.492	1,66	**0.485**	1093.824	0.007	0.106
2D:4D x Task x Sex x Response hand interaction	**12.681**	**1,66**	**p< 0.001**	**1093.824**	**0.161**	**0.939**

In order to break down the revealed four way interaction between right hand 2D:4D, task, sex, and response hand, analysis was split by sex and two separate three way ANOVAs were conducted in order to explore any main and/or interaction effects of digit ratio (low vs. high), sex (male vs. female), and response hand (left vs. right) on reaction times to the subitizing and control tasks considered separately. Analysis of subitizing response times revealed no significant main effect of 2D:4D, $F_{(1,66)} = 0.197$, MSe = 10579.829, $p = 0.659$, $\eta p^2 = 0.003$, $1\text{-}\beta = 0.069$, no significant sex x 2D:4D interaction, $F_{(1,66)} = 0.166$, MSe = 10579.829, $p = 0.685$, $\eta p^2 = 0.003$, $1\text{-}\beta = 0.069$, and no significant 2D:4D x response hand interaction, $F_{(1,66)} = 2.084$, MSe = 832.91, $p = 0.154$, $\eta p^2 = 0.01$, $1\text{-}\beta = 0.127$. A significant three way interaction between the factors 2D:4D x sex x response hand however was revealed, $F_{(1,66)} = 7.431$, MSe = 832.91, $p = 0.008$, $\eta p^2 = 0.101$, $1\text{-}\beta = 0.766$. As can be seen in figure 6a high 2D:4D (low PT) males displayed faster reaction times as compared to low 2D:4D (high PT) males for responses with both the left and right hands. High 2D:4D males also showed a more pronounced right hand advantage for the subitizing task than low 2D:4D males. A similar pattern of results however was not observed when considering female reaction times. As shown in figure 6b, while high 2D:4D females demonstrated faster subitizing reaction times than low 2D:4D females when using the left hand, low 2D:4D females displayed faster right hand response times in comparison to high 2D:4D females. In addition, in direct contrast to the male data, a stronger right hand advantage was observed for low 2D:4D females relative to high 2D:4D females.

No significant main effect of sex, $F_{(1,66)} = 0.004$, MSe = 10579.829, $p = 0.952$, $\eta p^2 = 0.00005$, $1\text{-}\beta = 0.05$, or sex x response hand interaction, $F_{(1,66)} = 2.084$, MSe = 832.91, $p = 0.154$, $\eta p^2 = 0.031$, $1\text{-}\beta = 0.296$, was observed. Similar to the overall 4-way ANOVA a significant main effect of response hand (right hand advantage) was revealed for the separate analysis of subitizing data, $F_{(1,66)} = 39.466$, MSe = 832.91, $p < 0.001$, $\eta p^2 = 0.374$, $1\text{-}\beta = 1$, see table 11.

a)



b)

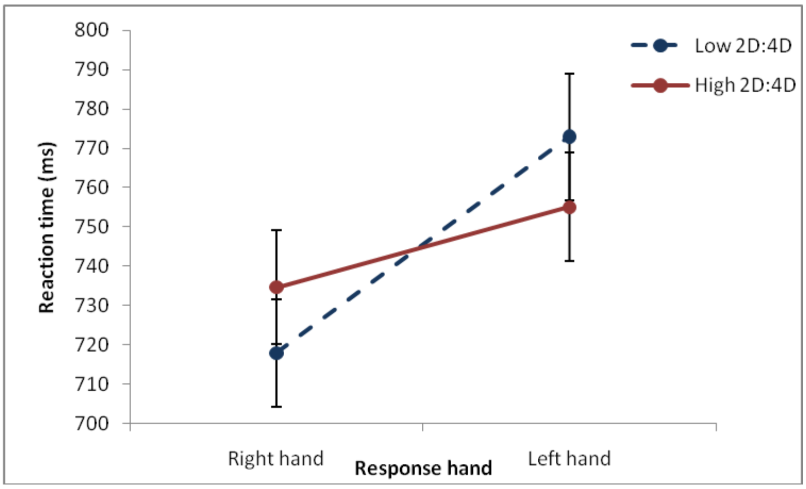


Figure 6 a & b. Mean subitizing reaction times in male (a) and female (b) low and high 2D:4D participants for responses with both the right and left hands.

In order to more closely explore the 3-way interaction between 2D:4D, response hand and sex for subitizing reaction times analysis was further broken down by sex and two separate 2-way ANOVAs were conducted to investigate any possible main effect of 2D:4D and response hand and their interactions on subitizing response times in males and females considered separately.

Neither males nor females showed an overall main effect of digit ratio (males - $F_{(1,32)} = 0.256$, MSe = 14564.37, $p = 0.617$, $\eta p^2 = 0.008$, $1\text{-}\beta = 0.078$; females - $F_{(1,34)} = 0.001$, MSe = 6829.674, $p = 0.975$, $\eta p^2 = 0.00003$, $1\text{-}\beta = 0.05$). Analysis of female data however did reveal a significant interaction effect between 2D:4D and response hand, $F_{(1,34)} = 6.735$, MSe = 799.6, $p = 0.014$, $\eta p^2 = 0.165$, $1\text{-}\beta = 0.713$. The interaction between 2D:4D and response hand was not found to be significant in male data analysed separately, $F_{(1,32)} = 1.697$, MSe = 868.302, $p = 0.202$, $\eta p^2 = 0.05$, $1\text{-}\beta = 0.244$. Both males and females showed a significant main effect of response hand (right hand advantage; males – $F_{(1,32)} = 10.909$, MSe = 868.302, $p = 0.002$, $\eta p^2 = 0.254$, $1\text{-}\beta = 0.893$; females – $F_{(1,34)} = 31.99$, MSe = 799.6, $p < 0.001$, $\eta p^2 = 0.485$, $1\text{-}\beta = 1$).

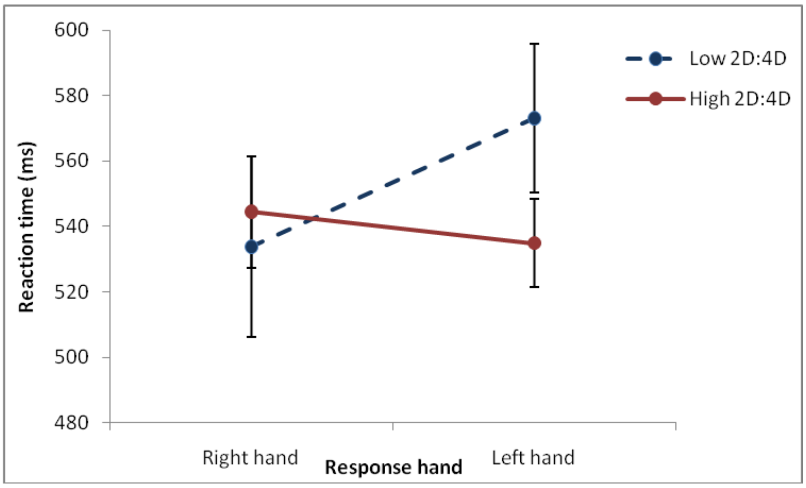
Post-hoc, Bonferroni corrected ($\alpha = 0.0125$) t-tests conducted in order to further explore the interaction between 2D:4D and response hand in females revealed a significant effect of response hand in low 2D:4D females, $t_{(17)} = 7.99l$ $p < 0.001$. No significant effects of response hand however was revealed in high 2D:4D females, $t_{(17)} = 1.786$; $p = 0.092$, and no significant effect of 2D:4D group was found in either right hand, $t_{(34)} = 0.836$; $p = 0.409$, or left hand , $t_{(34)} = 0.845$; $p = 0.404$, response times considered separately.

The separate analysis of performance on the control tasks also showed no significant main effect of 2D:4D on reaction times, $F_{(1,66)} = 1.381$, MSe = 9921.484, $p = 0.244$, $\eta p^2 = 0.02$, $1\text{-}\beta = 0.212$, and no significant 2-way interaction effects between 2D:4D group and response hand, $F_{(1,66)} = 1.89$, MSe = 1684.961, $p = 0.174$, $\eta p^2 = 0.028$, $1\text{-}\beta = 0.273$, or 2D:4D group and sex, $F_{(1,66)} = 0.129$, MSe = 9921.484, $p = 0.721$, $\eta p^2 = 0.002$, $1\text{-}\beta = 0.064$. Similar to the separate analysis of subitizing reaction times however, analysis of control reaction times also revealed a significant 3 way interaction between 2D:4D group, response hand and sex, $F_{(1,66)} = 4.585$, MSe = 1684.961, $p = 0.036$, $\eta p^2 = 0.065$, $1\text{-}\beta = 0.56$. The pattern of the interaction however was not similar to that observed for subitizing reaction times. As can be seen in figure 7a low and high 2D:4D males showed different response hand biases with low 2D:4D (high PT) males demonstrating a right hand advantage, while high 2D:4D (low PT) males displayed a left hand advantage. The degree of response hand differences in reaction times on the control task was greater in low 2D:4D males in comparison to high 2D:4D males.

A similar pattern was not identified in females whereby both low and high 2D:4D females showed a right hand advantage on the control task and the degree of response hand difference was greater in high 2D:4D females (low PT) relative to low 2D:4D females. Also in contrast to male data, high 2D:4D females displayed faster reaction times than low 2D:4D females for both right and left hand responses.

While high 2D:4D males however showed comparatively faster reaction times than low 2D:4D males for left hand responses, low 2D:4D males showed comparatively faster reaction times than high 2D:4D males for right hand responses. Separate analysis of control data also revealed a significant right hand advantage (see table 11) on the task, $F_{(1,66)} = 6.34$, MSe = 1684.91, $p = 0.014$, $\eta p^2 = 0.088$, $1\text{-}\beta = 0.699$. No significant main effect of sex, $F_{(1,66)} = 0.294$, MSe = 9921.484, $p = 0.59$, $\eta p^2 = 0.004$, $1\text{-}\beta = 0.083$, or response hand x sex interaction effect, $F_{(1,66)} = 0.139$, MSe = 1684.961, $p = 0.71$, $\eta p^2 = 0.002$, $1\text{-}\beta = 0.066$, was found for control reaction times.

a)



b)

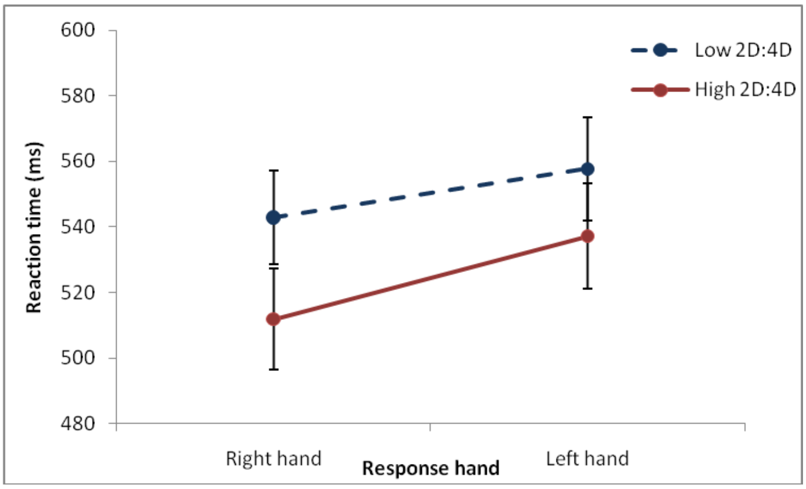


Figure 7 a & b. Mean control task reaction times in male (a) and female (b) low and high 2D:4D participants for responses with both the right and left hands.

The pattern of results responsible for the 3 way interaction between 2D:4D group, sex and response hand on the subitizing task is not comparable to the pattern of result responsible for the same 3 way interaction on the control task. As the purpose of the control task was to consider the possibility that any identified relationship between 2D:4D and performance on the subitizing task was not secondary to a more general relationship between 2D:4D and generic reaction time on a similar speeded response task, and the specific relationship between 2D:4D and generic reaction time is not of particular interest to the current study further analysis of the interaction between 2D:4D , sex and response hand on control task reaction times will not be reported here. The results of subsequent analysis of this interaction however can be viewed in appendix 3.

5.3.2.2 Analysis including left hand 2D:4D

An identical four-way ANOVA was conducted including the factor of left hand 2D:4D rather than right hand 2D:4D. Similar to data including right hand 2D:4D measures, no main effect of left hand 2D:4D was found. As can be seen in table 14, a significant 3 way interaction between the factors left hand 2D:4D, task and response hand was identified. No further significant interactions involving 2D:4D were observed, thus the four way interaction between right hand 2D:4D, task, sex and response hand was not replicated for left hand 2D:4D measures, see table 14. Results of the analysis including left hand 2D:4D data also revealed an overall main effect of task on reaction times, $F_{(1,66)} = 1374.122$, MSe = 2111.06, $p < 0.001$, $\eta p^2 = 0.954$, $1\text{-}\beta = 1$, and an overall main effect of response hand, $F_{(1, 66)} = 28.244$, MSe = 1453.009, $p < 0.001$, $\eta p^2 = 0.3$, $1\text{-}\beta = 0.999$, both effects were in a similar direction to those identified in the analysis including right hand 2D:4D measures.

Table 14

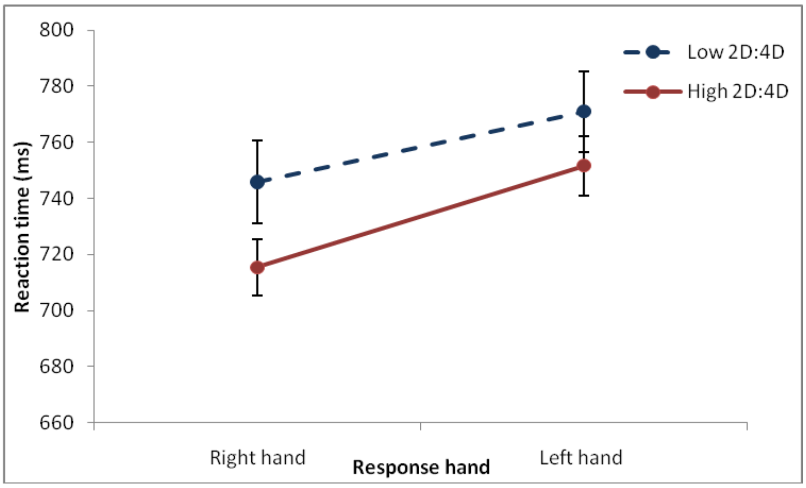
F, p, df, MSe, effects size (partial eta squared - ηp^2) and power ($1\text{-}\beta$) values relating to main and interaction effects for the factor of right hand 2D:4D (significant effect highlighted in bold).

Effect	**F value**	**df**	**p**	**MSe**	**ηp²**	**1-β**
Main effect 2D:4D	2.086	1,66	0.153	1805.308	0.031	0.296
2D:4D x Task interaction	0.103	1,66	0.75	2111.006	0.002	0.061
2D:4D x Sex interaction	0.036	1,66	0.85	18075.308	0.001	0.054
2D:4D x Response hand interaction	0.63	1,66	0.43	1453.009	0.009	0.122
2D:4D x Task x Sex interaction	0.406	1,66	0.526	2111.006	0.006	0.096
2D:4D x Response hand x Sex interaction	0.302	1,66	0.584	1453.009	0.005	0.084
2D:4D x Task x Response hand interaction	**5.14**	**1,66**	**0.027**	**1207.848**	**0.072**	**0.608**
2D:4D x Task x Sex x Response hand interaction	0.536	1,66	0.467	1207.848	0.008	0.111

With regard to the three way interaction between 2D:4D, task and response hand, as can be seen in figure 8, high 2D:4D participants showed faster reaction times than low 2D:4D participants on both the subitizing and control task for responses with both the left and right hands. For each task both high and low 2D:4D participants also showed a right hand advantage. On the subitizing task this right hand advantage was slightly more pronounced in high 2D:4D participants as compared to low 2D:4D participants. On the control task however the right hand advantage was comparatively more pronounced in low 2D:4D participants than high 2D:4D participants. In order to further investigate the significant interaction between left hand 2D:4D, task and response hand two separate 2-way ANOVAs were conducted to explore any main or interaction effects of 2D:4D and response hand on subitizing and control task reaction times considered independently. Separate analysis of subitizing and control task reaction times however revealed no significant main effect of left hand 2D:4D (subitizing – $F_{(1,68)} = 2.174$, MSe = 10004.271, $p = 0.145$, $\eta p^2 = 0.031$, $1\text{-}\beta = 0.307$; control – $F_{(1,68)} = 1.656$, MSe = 9661.146, $p = 0.202$, $\eta p^2 = 0.024$, $1\text{-}\beta = 0.245$) and no significant 2D:4D x response hand interaction effect (subitizing – $F_{(1,68)} = 1.185$, MSe =

981.791, $p = 0.28$, $\eta p^2 = 0.017$, $1\text{-}\beta = 0.189$; control – $F_{(1,68)} = 3.43$, $MSe = 1709.048$, $p = 0.068$, $\eta p^2 = 0.048$, $1\text{-}\beta = 0.447$) for either task. A significant main effect of response hand however (right hand advantage) was revealed for analysis of performance on both the subitizing ($F_{(1,68)} = 35.857$, $MSe = 918.791$, $p < 0.001$, $\eta p^2 = 0.345$, $1\text{-}\beta = 1$), and control ($F_{(1,68)} = 6.572$, $MSe = 918.791$, $p = 0.013$, $\eta p^2 = 0.088$, $1\text{-}\beta = 0.715$) tasks.

a)



b)

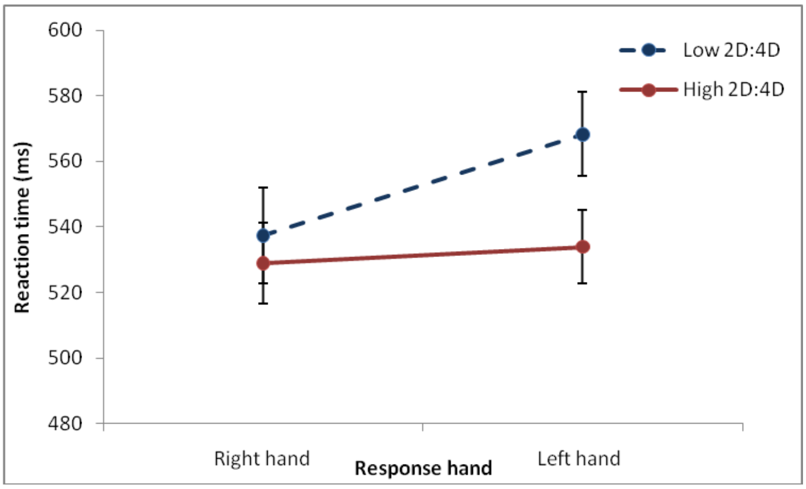


Figure 8 a & b. Mean subitizing (a) and control (b) task reaction times low and high 2D:4D participants for responses with both the right and left hands.

5.4 Discussion

The primary aim of the current study was to re-examine any potential relationship between 2D:4D, and the basic numerical process of subitizing, partially replicating the effect initially explored in experiment 2, but incorporating several methodological improvements.

In support of the established sex difference in 2D:4D (Phelps, 1952; George, 1930; Manning, Barley, et al. 2000) findings revealed significant sex differences in right hand 2D:4D measures, such that males demonstrated significantly lower 2D:4D ratios assumed to reflect higher PT in comparison to females. In accord with experiment 2 however, while in the anticipated direction, sex differences in left hand 2D:4D measures were not found to be significant.

No main effects of either right or left hand 2D:4D were revealed suggesting that 2D:4D is not associated with overall performance on both tasks. The results also failed to identify any significant correlations between 2D:4D and subitizing reaction times or any significant interaction between either right or left hand 2D:4D and task. Overall therefore the results suggest the 2D:4D does not appear to directly relate to the primitive ability to subitize. According to this interpretation any associations that may exist between speed of access to subitizing and higher mathematical ability are thus unlikely to directly account for any potential relationship that may exist between aspects of numerical and mathematical competency and PT exposure. As subitizing however represents only one aspect of core mathematical processing it is impossible to rule out a potential influence of PT on other aspects of core magnitude processing, which may also form a foundation for higher mathematical competence.

The above interpretation may also be premature given that, similar to experiment 2, power calculations computed for both correlation and ANOVA analysis suggest low power in the current experiment. While a number of the effects sizes for both sets of analyses were also extremely low, i.e. well below what would be classified as a small effect according to Cohen’s reported conventions (Cohen, 1988), some of the reported effect sizes were within a small-medium range. What’s more, where small to moderate effects were observed the findings are in line with the direction of the non-significant small to moderate effects reported in experiment 2 and the finding reported by Fink et al. (2006) and Brosnan (2006) such that low 2D:4D was associated with improved performance in males and poorer performance in females. The sample size in the present study however was below that which would be required in order to achieve a small to moderate effect size for either correlation or ANOVA analysis according to the prospective power calculations described in experiment 2 (see section 4.4, p. 92-93). Again therefore it is important to recognise that null results in the current study may be related to lack of power.

While no main effects of either right or left hand 2D:4D on performance on the subitizing task were found, a significant 4-way interaction was revealed between the factors of right hand 2D:4D, sex, task, and response hand. A significant 3-way interaction was also revealed between the factors of left hand 2D:4D, task and response hand. Subsequent analysis of the 4-way interaction revealed a three way interaction between right hand 2D:4D, sex, and response hand for performance on the subitizing task. Closer inspection of the result suggested that the factors of right hand 2D:4D and response hand may interact differently depending upon sex for performance on a subitizing task. A significant interaction between right hand 2D:4D and response hand was observed in females. While both low and high 2D:4D females showed a right hand advantage for the task, the effect of response hand was only significant for low 2D:4D (high PT) females, suggesting a stronger right hand advantage in low 2D:4D females as compared to high 2D:4D (low PT) females. In males both low and high 2D:4D participants also showed a right hand advantage, contrary to females however, this advantage was slightly stronger in high 2D:4D males. The interaction between right hand 2D:4D and response hand however was not found to be significant in analysis of male data. A three way interaction was also revealed between right hand 2D:4D, sex and response hand on the control task. Once again the factors of right hand 2D:4D and response hand appeared to interact differently for males and females, crucially however a dissimilar pattern of result was observed to that identified for subitizing task performance. Contrary to the results revealed for subitizing the interaction between right hand 2D:4D and response hand was significant for males but absent in females.

With reference to the three way interaction between left hand 2D:4D task and response hand, visual inspection of the results implies different patterns of interaction between left hand 2D:4D and response hand for the subitizing vs. the control task. Subsequent analysis however did not reveal a significant interaction between these two factors for either task.

As hand preference is perhaps the most distinct example of behavioural lateralization, one possible interpretation is that the findings of the current experiment may indicate a potential role for PT on lateralization for the process of subitizing. In support of this interpretation evidence suggests that right vs. left handed individuals may, demonstrate different patterns of hemispheric specialization (Levy & Reid, 1978; Pujol et al., 1999; Knecht et al., 2000). In addition there is evidence that extent of hand preference may be associated with the degree of functional lateralization (Dassonville, et al., 1997; Isaacs et al., 2006). According to such evidence, it seems plausible to suggest that the findings may implicate a possible role for PT on lateralization for basic numerical processing, with differential effects potentially in operation for males vs. females. To the extent that degree of manual asymmetry may reflect patterns of cerebral activation, the significant effect of response hand in low right hand 2D:4D females in the absence of an effect for high right hand 2D:4D females may implicate higher levels of PT exposure to correspond to greater lateralization for subitizing, with the opposite trend implied, although not revealed to be significant, for males.

It is important to recognise that the results of the current experiment also imply a significant relationship between 2D:4D and lateralization for control task performance, suggesting that the influence of PT may be generally related to lateralization during similar reaction times tasks. In the current study however the relationship between 2D:4D and lateralization was found to be different for control vs. subitizing task performance. Findings specific to subitizing therefore cannot easily be explained in terms of a generic effect of PT on patterns of lateralization for any speeded response task.

The potential link between 2D:4D and patterns of lateralization in the present study is particularly interesting given that findings from the previous study suggested differences in patterns of association between 2D:4D and visual field presentation for the process of subitizing vs. counting. Both studies therefore imply a possible relationship between 2D:4D and lateralization for subitizing task performance in contrast to performance on a different yet procedurally comparable speeded response task.

Crucially however, while analysis of both right and left hand 2D:4D data in the current study implied a possible association between 2D:4D, response hand and performance on the subitizing and control task it is unclear why the two analyses failed to produce more comparable results. The failure to replicate the significant 4-way interaction between right hand 2D:4D, sex, task and response hand for analysis of left hand data is partially in line with literature that generally reports the relationship between 2D:4D and psychological/cognitive factors to be stronger (or only present) for the right hand (e.g. Williams et al., 2000; Brown et al., 2002; Csathó et al., 2003; Lutchmaya et al., 2004). Previous research however does not provide any explanation for why the 3 way interaction between 2D:4D, task and response hand revealed for analysis left hand data was absent for analysis of right hand data.

In addition to problems regarding interpretation, the current study also possessed a number of important methodological limitations which should be identified. Firstly, reaction times to the colour blue on the control task were found to be significantly faster that those to the colour red and the colour yellow. While it is possible that results could reflect general differences in reaction times to different colours, given that the response button for the colour blue was positioned in the centre of the response box (blue) it seems more plausible that differences may reflect positional effects of response keys on reaction time.

More importantly, consideration of subitizing reaction time also supports the possibility that the position of the response keys may present an important confound in the data. As described previously (see chapter 4), subitizing response times are typically reported in the region of approximately 500-800ms (Trick & Pylyshyn, 1994) with the average subitizing RT x set size slope gradient documented by Dehaene & Cohen (1994) to be approximately 50ms for 2-3 items, and 100-200ms for 3-4 items, although some authors report lower slope gradients (e.g. Trick & Pylyshyn, 1994; Wender & Rothkegel, 2000). In the current study however average response time to the enumeration 2 and 3 dots was almost identical, the average reaction time for 3 dots was actually just below the average reaction time observed for 2 dots, although the difference was minimal. It is highly possible then that reaction times to the quantity 3 were facilitated by the fact that the necessary response button was positioned in the centre of the response box. It is important therefore that caution should be observed when interpreting the nature of behavioural measures for subitizing.

Secondly, even with brief presentation, while mean reaction times from the present (and previous) experiment are within the range of reported reaction times in the literature (500 – 800ms) they are nevertheless at the very top end of the range (mean = 758ms). As the procedure followed in the current (and previous) study is very similar to that used in the large majority of research investigating subitizing, the reason for this is unclear. One potential difference is that some authors also employ a backwards mask in order to control for the potential effects of retinal after-images on reaction times. It is possible in the present experiment that some participants may still potentially be employing a counting as opposed to subitizing strategy on the basis of these retinal after-images.

Finally while the current study did not specifically aim to address associations between 2D:4D and patterns of lateralization for the process of subitizing a number of interesting findings were revealed for a potential association between 2D:4D and response hand. Results relating to response hand have in turn been discussed with reference to a potential link between 2D:4D and lateralization on the subitizing task relative to the control task. As handedness however is thought to be influenced by a variety of factors, including maternal handedness and family history of sinistrality (Annett, 1998; 1999), and degree of hand preferences may vary according to the task or activity being observed (Annett, 1985; Steenhuis & Bryden, 1989; Gilbert & Wysocki, 1992), the measure does not constitute a particularly reliable technique for the assessment of hemispheric lateralization during cognitive processing. Unsurprisingly therefore, hand preferences are not typically assessed in order to elucidate patterns of lateralization during task performance.

The inclusion of lateralized stimuli presentations, as utilised in experiment 2, constitutes a far more conventional and widely recognised method for the exploration of hemispheric asymmetry of function. Furthermore, the level of PT exposure has been hypothesised to influence developing patterns of left vs. right hand preferences (Geschwind & Galaburda, 1987; Witelson, 1991; Witelson & Nowakowski, 1991). Of particular interest to the present study, evidence utilising 2D:4D as a marker of PT has reported low 2D:4D (high PT) to be associated with a lower degree of right handedness in right handed participants (Fink et al., 2004). While all participants in the current experiment were right handed according to self reported writing hand, individual variation in degree of right handedness was not controlled for. Possible associations between 2D:4D and degree of handedness therefore may have potentially biased the results.

Future research should aim to usefully address the above identified methodological limitations. In addition it would be useful to also explore possible relationships between 2D:4D and alternative measures of core numerical processing. As noted previously, subitizing is just one aspect of basic numerical processing, which is thought to serve as a foundation for more sophisticated numerical concepts. Feigenson et al. (2004) identified two core numerical systems that may form the basis of basic numerical intuition: (1) a system for the precise representation of small quantities (subitizing) and (2) a system for representing large approximate magnitudes. Beyond subitizing, further research might examine potential links between the approximate system of basic numerical representation, 2D:4D, sex, and, given the results of the present and previous study, specific patterns of lateralization.

In summary, in contrast to previous research but similar to findings described in experiment 2, the present study identified no simple pattern of associations between 2D:4D and performance for the basic numerical task of subitizing. Significant interactions however were identified which imply different patterns of association between 2D:4D and response hand preference for subitizing vs. control task performance. Analysis of right hand 2D:4D data revealed a significant 3 way interaction effect between 2D:4D, sex and response hand on subitizing reaction times, further implying a potential influence of sex on the association between 2D:4D and response hand preferences. Female participants with low 2D:4D ratios showed a significant right hand advantage on the subitizing task while no significant response hand differences were observed for females with high 2D:4D ratios.

One possible interpretation is that behavioural lateralization manifested in the degree of a particular hand advantage, may reflect patterns of cerebral lateralization on a particular task. According to this speculative interpretation, then, the results imply that higher levels of exposure to PT in females results in a greater degree of lateralization for subitizing. Such findings are intriguing given that results from the previous chapter may also imply process dependent patterns of lateralization during counting vs. subitizing. Both studies however did possess a number of methodological limitations. Further investigation of these potential relationships is required in order to confirm possible effects.

Chapter 6

Experiment 4: 2D:4D and Basic Numerical Ability in Adults.

6.1 Introduction

While experiments 2 and 3 did not find any evidence for a direct relationship between subitizing and 2D:4D, they did suggest a potential link between 2D:4D and the extent of lateralization during subitizing task performance relative to performance on a comparable speeded response task. In experiment 2 different patterns of association between 2D:4D group (low vs. high) and visual field presentation for reaction times to the process of subitizing vs. counting were identified. No significant main effect of 2D:4D or 2D:4D x visual field interaction effect however was revealed for the process of subitizing or counting analysed separately.

After controlling for certain methodological factors, experiment 3 revealed a significant association between 2D:4D, sex and response hand preferences during subitizing. In females a significant association between 2D:4D and response hand was observed for subitizing reaction times with low 2D:4D females demonstrating a significant right hand advantage for the task. While high 2D:4D females also displayed a right hand advantage for subitizing reaction time, this advantage was not found to be significant. The opposite pattern of results was observed in males, with high 2D:4D males showing a comparably stronger right hand advantage for subitizing reaction times than low 2D:4D males; the interaction between 2D:4D and response hand in males however was not significant. While a significant interaction between 2D:4D, sex and response hand was also indentified for the control task, the pattern of results regarding control task reaction times was different to that revealed for subitizing reaction times. Specific findings relating to subitizing reaction times therefore are unlikely to be due to a general effect of prenatal testosterone (PT) upon tasks requiring speeded response. To the extent that degree of handedness in response to a particular task may reflect hemispheric asymmetry for a process, the results suggested that high foetal T in females may be associated with a greater degree of lateralization for subitizing. Furthermore, the overall three-way interaction between 2D:4D sex and response hand suggests that the effect of PT testosterone exposure on lateralization during subitizing may be different for males. This revealed dissociation between male and female findings in experiment 3 is similar to previous evidence suggesting that exposure to PT may differentially affect numerical and/or mathematical performance in men and women (Brosnan, 2008; Finegan et al., 1992; Fink et al., 2006) and may have important implications for hypotheses concerning the influence of PT on sex differences in numerical and mathematical aptitude.

It is vital to acknowledge however that as handedness is believed to be affected by multiple factors, including PT itself (Fink et al., 2004; Manning et al., 2000), preferences and/or degree of hand preference does not constitute a rigorous or reliable technique for the assessment of hemispheric specialization. Furthermore, experiment 3 failed to control for individual differences in degree of handedness preference across the sample. Such variation therefore may have potentially confounded the results. It should also be recognised that the findings of experiment 3 showed possible stimulus-response compatibility effects on reaction times. More specifically, the fastest average response times (on both the subitizing and control task) were observed in the recognition of the stimuli for which the necessary response button was positioned in the centre of the three key response box (keys were positioned in a straight line).

Given the novelty of the research and the consistency of revealed effects in experiments two and three with regard to a potential link between 2D:4D and lateralization for subitizing, the current study sought to provide further understanding of the putative relationships between 2D:4D and subitizing addressing previous methodological issues. In addition, the present experiment aimed to utilise the general procedure developed over the previous two experiments in order to also investigate possible relationships between 2D:4D and other 'core' and basic numerical tasks.

6.2 The current study

6.2.1 Subitizing

In order to re-investigate relations between subitizing and 2D:4D the same tasks (both subitizing and control) as those adopted in experiment 3 were employed with the following modifications; firstly, given the implied importance of lateralization in regard to potential relationships between 2D:4D, stimuli was presented to either the left or right visual field as opposed to centrally in order to assess hemispheric specialization in a more rigorous and conventional manner. Secondly, in order to avoid confounds relating to hand of response and response key positioning a voice key was used to measure reaction times, with participant responses recorded so as to also assess accuracy. Thirdly, a standardized measure of hand preference was employed in order to confirm self reported handedness. Finally, a backward mask was used to eliminate the possible effects of retinal afterimages on reaction times of enumeration. The decision to include this final modification was based on the observation that even with brief presentation, mean RTs in experiment 3 were relatively high when considered in context of those previously reported in prior literature (see experiment 3 discussion). One possible explanation is that some participants may still have been employing a counting, as opposed to a subitizing strategy on the basis of retinal after images, thus the use of a mask following stimulus display was deemed appropriate in order to counteract this possible confound in the current study.

6.2.2 Counting

In addition to exploring potential relationships between 2D:4D and subitizing, experiment 2 also addressed any potential associations between 2D:4D and the process of conventional counting. The results revealed a three way interaction between process (subitizing vs. counting), 2D:4D and visual field presentation on reaction times. Again however, due to methodological limitations, clear interpretations and conclusions regarding either process were impossible.

Experiment 2 utilised one common task in order to assess both subitizing and counting. The findings demonstrated however the importance of separate response deadlines when evaluating these two separate processes. Results from experiment 2 suggested that in the allocation of an unlimited response deadline participants are likely to have adopted a counting as opposed to subitizing strategy for the enumeration of quantities < 4. Evidence from Pasini and Tessari (2001) however suggests that the use of a fixed response time (as high as 3 seconds) may induce an estimation as opposed to counting strategy for quantities > 4. An assessment of counting therefore was not incorporated into experiment 3 so that specific focus could be given to potential relationships between 2D:4D and subitizing utilising suitable response deadlines.

As identified in chapter 4, despite serving similar purposes, behavioural, clinical and neuro-imaging studies suggest that the processes of subitizing and counting are distinct (Butterworth, 1999; Chi & Klahr, 1975; Cipolotti et al., 1991; Dehaene & Cohen, 1994; Mandler & Shebo, 1982; Nan et al., 2005; Piazza et al., 2002; 2003; Sathian et al., 1999; Trick & Pylyshyn, 1993, 1994). Importantly counting is not typically thought to be a 'core' numerical skill. Behavioural evidence has also reported that the two processes may demonstrate different patterns of hemispheric specialization, with Pasini & Tessari (2001) reporting a right hemispheric bias for subitizing and a left hemispheric bias for counting. Thus, as identified in experiment 2, any potential relationships between 2D:4D and counting provides an interesting source of comparison/contrast to possible associations between 2D:4D and subitizing. Utilising an entirely separate task to that adopted in order to assess subitizing therefore, the current study also attempted to re-explore any potential relationships between 2D:4D and counting. The RT time task adopted in order to investigate counting was similar to that used to investigate subitizing, differing only in regard to the stimuli (quantities 2, 3, and 4 for subitizing, and 6, 7, and 8 for counting), the stimuli display and response time (with a longer response presentation time and deadline adopted for counting), and the use of a backwards mask (not included in the counting task). Similar to the subitizing task, given previous links with lateralization, presentation of counting stimuli was lateralized. Also similar to subitizing, a comparable control task was also employed.

6.2.3 Number Comparison

As highlighted in previous chapters, subitizing constitutes just one aspect of 'core' numerical processing. The ability to approximately represent and compare large magnitudes has also been identified as a 'core' numerical skill, present in infants and animals (e.g. Barth et al., 2004; Brannon & Terrace, 1998; Hauser et al., 2003; Lipton & Spelke, 2003; Xu & Spelke, 2000) and thought to serve as a foundation for more sophisticated numerical concepts (Durant et al. 2005; Llanderl et al., 2004). Similar to subitizing therefore any potential link between 2D:4D and this second 'core' system of numerical processing may potentially underlie possible relationships between aspects of developing and higher mathematical and numerical abilities, in a way that is unlikely to be secondary to alternative cognitive skills.

As reviewed in chapter 3, this basic system for the representation of large approximate magnitudes is reflected in the ability, present in infants, non-human animals and adults to closely approximate quantities and/or discriminate the larger of two arrays, without counting and independent of stimulus modality. In all three groups this approximate representation of magnitude is subject to similar limitations, namely, the distance effect in which a systematic monotonous decrease in numerical discrimination performance occurs as numerical distance between numbers decreases and, the size effect in which the accuracy of numerical approximation and discrimination for equal numerical distances decreases with increasing number size (Dehaene, 2000).

The present study also attempted to explore relationships between PT exposure and this second 'core' system of number using a RT number comparison task based broadly on the design developed in order to assess subitizing. Participants were asked to identify the larger of two-arrays, and response time was measured using a voice key (vocal responses were recorded in order to assess accuracy). Given the revealed importance of lateralization with regard to revealed relationships between 2D:4D and subitizing, stimuli for the number comparison task was presented to either the left or right visual fields in order to also explore potential 2D:4D effects as a function of hemispheric specialization for the task. Again, a control task was also included, comparable to the experimental task in all aspects excluding those relating to numerosity.

6.2.4 SNARC effect

The mental representation of our apparently innate, phylogenetically developed ability to approximate and discriminate large quantities is hypothesised to be analogous to a logarithmically compressed, analogue 'mental number line' (Dehaene et al., 1992). Support for this hypothesis can be derived from evidence for an association between speed of responses to numerical magnitude and the spatial format of required responses. A large body of evidence now suggests that faster reaction times result when responding to small quantities/numbers with the left as opposed to right hand and when responding to large quantities/numbers with the right as opposed to left hand. This association is termed the Spatial Numerical Association of Response Codes (SNARC) effect and, as discussed in chapter 3, has recently been reported to be associated with 2D:4D. Intriguingly, Bull and Benson (2006) observed a stronger SNARC effect in low 2D:4D (high PT) participants as compared to high 2D:4D (low PT) participants, potentially suggesting a stronger mental representation of numerical magnitude along a spatially orientated mental number line in this group. These authors however failed to control for generic reaction time. As SNARC-style effects have also be observed for alternative categories of ordinal stimuli such as letters of the alphabet, days of the week and months of the year (Gevers et al., 2003; 2004), it is possible that the results obtained by Bull and Benson (2006) simply reflect a relationship between exposure to PT and a general reaction times for the automatic representation of spatially organised ordinal sequences. The present study thus attempted to replicate the findings of Bull and Benson (2006), using an identical procedure, with the additional element of a comparable control task so to assess general reaction time to an alternative series of spatially composed ordinal stimuli.

6.2.5 Hypotheses

Based on evidence from experiments 2 and 3 it is hypothesised that the results will reveal a relationship between 2D:4D and visual field of stimulus presentation for performance on the subitizing and counting tasks. It is anticipated that this relationship will be independent of any general association between 2D:4D and visual field that may exist for performance on the control task. On the basis of this prediction is thus hypothesised that significant three way interactions will be identified between 2D:4D, visual field and task (numerical vs. control) for both subitizing and counting task performance. In light of the findings of experiment 3 it is further hypothesised that any relationship between the factors of 2D:4D, task and visual field of stimulus presentation may also interact with sex such that any revealed 2D:4D group differences in visual field preferences during subitizing task performance relative to the control task performance are different in males and females or only significant in one sex. Based on the results of experiment 3 it is predicted that that low 2D:4D in females will be related to increased lateralization (as reflected by greater visual field differences) for the subitizing task.

On the basis of previous research evidence for potential relationships between correlates of PT and numerical and mathematical task performance (Brosnan, 2008; Finegan, et al., 1992; Fink et al., 2006; Kempel et al., 2005; Luxen & Buunk, 2005, see chapter 1 for review) it is also hypothesised that an association may exist between 2D:4D and number comparison task performance and 2D:4D and the SNARC effect. Once again it is anticipated that any identified relationships between 2D:4D and performance on either tasks will be distinct from any general relationships which might exist between 2D:4D and performance on non-numerical yet procedurally comparable control tasks. On the basis of findings reported by Bull and Benson (2006) it is predicted that participants with lower 2D:4D (higher PT) will show a more pronounced spatial representation of magnitude reflected in a stronger SNARC effect in comparison to high 2D:4D participants. Given the results of experiment 3 it is further predicted that any effect of 2D:4D on number comparison relative to control may further interact with the visual field of stimulus presentation and that any effect of 2D:4D on both number comparison and SNARC task performance relative to control may interact with sex. Based on previous findings it is hypothesised that any revealed relationships between 2D:4D, number comparison and SNARC task performance, will be significant in one sex only, or will reflect a different pattern of results between the two sexes. In line with the procedure adopted in experiments 2 and 3 raw 2D:4D data will be used to categorise participants into low vs. high 2D:4D groups in order to facilitate the exploration of potential complex interactions with the factors of task (experimental vs. control), sex and visual field presentation.

6.3 Method

6.3.1 Design

The study employed a 2 x 2 x 2 x 2 mixed measures, quasi-experimental design. Separate numerical and control tasks were employed for each aspect of numerical performance being considered (i.e. subitizing, counting number comparison and the SNARC effect). The study aimed to evaluate any main or interaction effects the factors; 2D:4D (low vs. high), sex (male vs. female) and task (numerical vs. control) on performance. The factor of visual field of stimulus presentation was also included when considering subitizing, counting and number comparison performance. Similar to experiments 1, 2 and 3, basic associations between right and left hand 2D:4D and numerical task performance were also explored using correlations.

6.3.2 Participants

A total of 80 participants were recruited from the student population of Northumbria University via e-mail and poster advertisement. All participants gave their full written informed consent to partake in the experiment and were paid £7 for their participation. Degree and direction of handedness was assessed according to the Edinburgh Handedness Inventory (EHI, see Appendix 12) (Oldfield, 1971). All participants reported normal or corrected to normal visual acuity, no present or previous injury to the second or fourth fingers of either hand, and no known hormonal abnormalities. Employing the exclusion criteria of experiment 2 (i.e. excluding all participants not of the majority handedness, sexual orientation and ethnicity) led to a final sample of 70 (35 males and 34 females) right handed, white British, heterosexual participants (ethnicity and sexual orientation assessed according to self report).

Unfortunately however, due to technical difficulties, not all participants completed each task resulting in varied sample sizes across task data sets. Overall, 63 participants, 33 males (mean age = 20.52, SD = 1.6) and 30 females (mean age = 20.83, SD = 3.54) completed subitizing and subitizing control tasks; 61 participants, 32 males (mean age =20.72, SD =1.78) and 29 females (mean age = 21.03, SD = 3.56) completed counting and counting control tasks; 66 participants, 33 males (mean age =20.67, SD =1.8) and 33 females (mean age = 20.91, SD = 3.39), completed number comparison and number comparison control tasks; and 67 participants, 35 males (mean age = 20.57, SD = 1.79) and 32 females (mean age = 20.88, SD = 3.41) completed the SNARC and SNARC control task. As the complete deletion of a single participant's data due to missing information on only one task would have resulted in a notable reduction in sample size, in order to maintain power, analysis of the effect of 2D:4D on performance was conducted entirely separately for each subset of main and control tasks.

6.3.3 Reaction time tasks

All tasks were created and administered using the experiment generator package Direct RT (Empirisoft software) and conducted on a Toshiba Tecra M1, Intel Centrino processing laptop with a 14.1" SXGA+ screen. In an attempt to aid concentration, half way through each task computerised programmes paused in order to offer participants a short break (approximately 5 minutes maximum) before continuing. Visual stimuli for each task were again developed using Microsoft ‘Paint’. Instructions for all tasks emphasised both speed and accuracy.

6.3.3.1 Subitizing and subitizing control

For both subitizing and counting stimuli, dot patterns were arranged within an invisible parameter of a 6×6cm square with the internal side of each square positioned 4.5cm either left or right of central fixation so that each dot pattern (dots = 0.8º visual angle) was viewed at an eccentricity varying from 4.5˚ to 9.5˚ of visual angle. Minimum and maximum distances between items were thus 0.4 and 4.4 respectively in the horizontal and vertical directions. For both tasks eight sets of different dot patterns were randomly generated for each quantity, and presented in a pseudo-random order such that each quantity was presented 16 times to each visual field (32 times in total). Subitizing and counting stimuli were developed using a random number generator in order to place elements into one of 36 numbered grid sections within the 6×6cm square.

Subitizing task stimuli consisted of the quantities two, three, and four presented in the form of black dots on a white background to either the left or right of the screen (see figure 9a). Prior to every trial, a central fixation (X) appeared on the screen for a duration of 1000ms. Stimuli were presented 1000ms after fixation for a total of 100ms before being masked by a display demonstrating two boxes containing oblique lines at left and right stimuli presentation areas (see figure 9b). Participants were given 1500ms in which to state out loud, as quickly and accurately as possible, the number of dots presented. The masking display remained for the full 1500ms response period or until participant response. Responses were measured to the nearest millisecond via a voice key triggered by response onset. Responses were also recorded in order to assess accuracy. Following participant response or the termination of the response deadline, inter-trial-interval was randomly set between 1000ms and 1500ms. As a control for general reaction time, all participants also performed a simple colour recognition task in which a 6cm × 6cm coloured square (either red, yellow, or blue) was presented to either the left of right of the screen using identical parameters of duration and presentation as the subitizing task (see figure 9d). Participants were instructed to state as quickly and as accurately as possible the name of the colour presented. Both the subitizing and subitizing control tasks began with a practice block of 12 trials, 6 leftward presentations, 6 rightward presentations, of randomly selected arrays/colours.

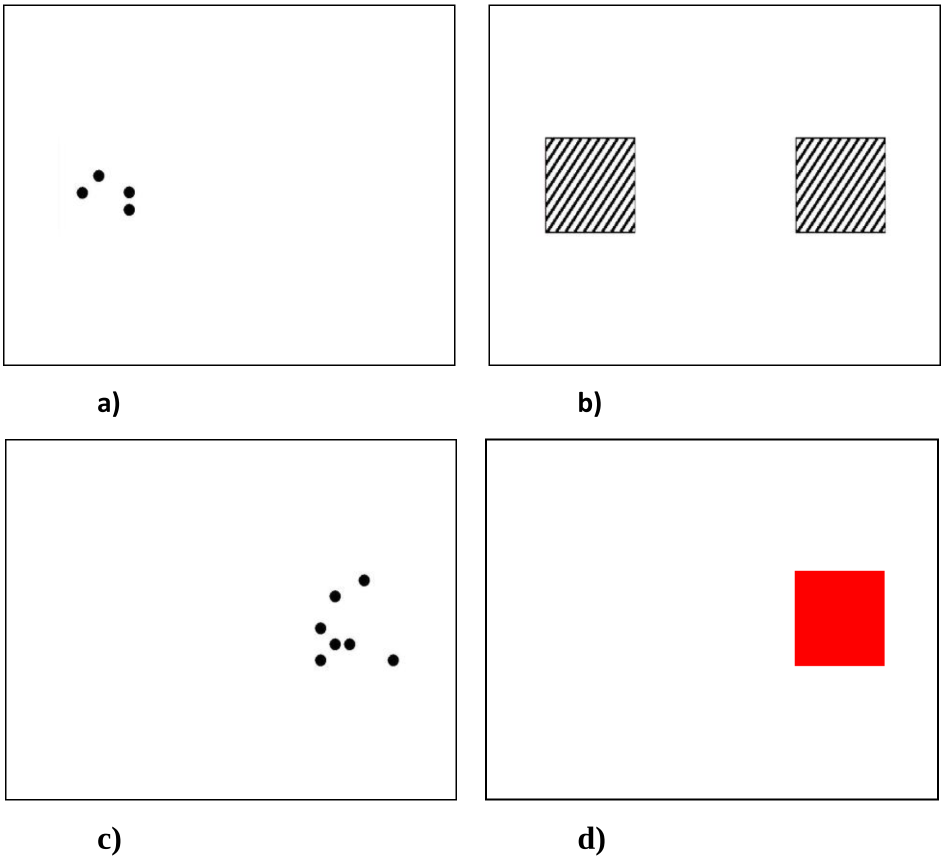


Figure 9. Examples of; a) subitizing stimuli presented to the left visual field, b) the masking display employed during the subitizing and subitizing control task procedure, c) counting stimuli presented to the right visual field and d) control task stimuli (both subitizing and counting task) presented to the right visual field.

6.3.3.2 Counting and counting control tasks

Counting task stimuli consisted of the quantities six, seven and eight presented in the form of black dots on a white background (see figure 9c). Besides the difference in stimuli, parameters of presentation were identical to that described for the subitizing tasks with the exception that stimuli remained onscreen until participant response thus a backward mask was not employed following stimulus presentation. Again a comparable control of general reaction time was adopted in which participants were asked to perform a simple colour recognition task (see figure 9d). The procedure for this task was identical to the control subitizing task but with the same exception that stimuli remained onscreen until participant response and thus a backward mask was not adopted.

6.3.3.3 Number comparison and number comparison control procedures.

For the number comparison task participants were tested with arrays of 10 – 30 red and blue dots presented too briefly for counting. Dots of two sizes, either 2 mm or 5 mm in diameter, appeared within virtual rectangular enclosures of two sizes, either 9 x 6cm² or 7 x 5 cm² (see figure11a). Participants were presented with 2 arrays of dots, one consisting entirely of blue dots and the other entirely of red dots, positioned above and below one another extending 7º horizontally and 5º vertically either left or right, above or below fixation (measured from fixation to the centre of virtual rectangular enclosures). Numerosities of the sets differed by ratios of either 0.57, 0.67 or 0.8 (4:5, 4:6, 4:7). For each ratio 4 possible number combinations were utilised. Stimuli were presented in a pseudo-random order such that comparisons for each ratio were presented 16 times to each visual field (32 times in total). For each of these 16 sets of trials per visual field, 50% demonstrated a negative correlation between number and; dot size, total contour length, summed dot area, and density (and thus a positive correlation between number and the size of the virtual rectangular enclosing), while the remaining 50% showed the reverse correlation. Any response patterns based on continuous variables as opposed to numerical value therefore would appear at chance overall. The chosen ratios of comparison and the measures taken in order to control for the possible influence of surface area and density on number comparison judgements were based on the procedure utilised by Barth et al. (2005) in order to explore numerical comparison in children.

For each half of these 16 sets, per visual field, the colour and position of the smallest array was counterbalanced across trials, so that; the smallest quantity constituted the red array and appeared above the largest quantity on 25% of trials, the smallest quantity constituted the red array and appeared below the largest quantity on 25% of trials, the smallest quantity constituted the blue array and appeared above the largest quantity on 25% of trials, and the smallest quantity constituted the blue array and appeared below the largest quantity on 25% of trials. Prior to every trial, a central fixation (X) appeared on the screen for a duration of 1000ms. Stimuli were presented 1000ms after fixation for 2500ms. Participants were instructed to state out loud the colour of dots in the array with the greatest numerosity. Responses were measured to the nearest millisecond via a voice key triggered by response onset. Responses were also recorded in order to assess accuracy. Inter-trial interval was randomly set between 1000ms -1500ms and began following either participant response or termination of the response deadline.

In order to control for general reaction time on a similar speeded response task a comparable size comparison task was also included. For this task two differently sized coloured squares (one red and one blue) were presented to either the left or right of the screen and participants were required to identify the colour of the largest square (see figure 11b). Similar to the number comparison task squares were presented above and below each other extending 7º horizontally and 5º vertically either left or right, above or below fixation (measured from fixation to the centre of the square). The surface areas of the two coloured squares differed by the same ratios utilised for numerosity comparison in the number comparison task, namely 0.57, 0.67, 0.8. For each ratio the surface area values were the same as the values adopted for number comparison, for example, one of the combinations adopted during the number comparison task for the ratio 0.57 was 12 dots vs. 21 dots, thus in one of the combinations for the 0.57 ratio size comparison on the control task one square had a surface area of $12cm^2$ and one of $21cm^2$.

All other parameters of presentation and duration similarly reflected those employed on the number comparison task. Notably, in line with the number comparison task the colour and position of the smallest square was counterbalanced across trials so that, on 25 % of trials, equally distributed across presentation in both visual fields; the red square displayed the smallest surface area and appeared above the blue square, the red square displayed the smallest surface area and appeared below the blue square, the red square displayed the largest surface area and appeared above the blue square, and the red square displayed the largest surface area and appeared below the blue square. Both the number comparison and size comparison control tasks were preceded by a practice block of 12 (6 rightward presentations and 6 leftward presentations) randomly selected trials.

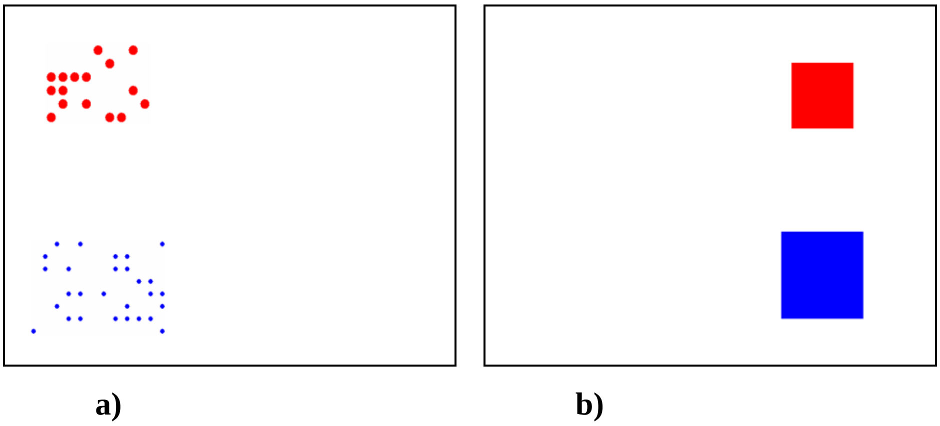


Figure 10: Examples of the; a) number comparison task stimuli presented to the left visual field, b) number comparison control task stimuli presented to the right visual field.

6.3.3.4 SNARC and SNARC control tasks

A similar procedure to that adopted by Bull and Benson (2006) was utilised in order to assess the presence of the SNARC effect. Digits 1-9 (in 72 point Times New Roman font) were presented to the centre of the computer screen for a maximum of 1300ms during which participants were required to identify whether the displayed integer was either an odd or even number. Prior to every trial, a central fixation (X) appeared on the screen for a duration of 1000ms followed by stimuli presentation 1000ms after fixation. Half of the participants responded by pressing a leftward key (Z) for odd numbers with the left hand and a rightward key (/) for even digits with the right hand, while the remaining participants began with the opposite key assignment (left-even, right-odd). Response key assignment was then reversed half way through the task. Each digit was presented 7 times in each order of response hand block (14 times in total). Participants also completed 18 practice trials of randomly selected digits per block in order to become acclimated to the procedure and required response. Inter-trial interval was randomly set between 1000msec and 1500msec and began following participant response or termination of response deadline.

In an extension of the Bull and Benson (2006) paradigm, the current study also incorporated a comparable reaction time task employing non-numerical stimuli with a similar ordinal organisation to numbers, so to control for the possibility that any potential relationships between 2D:4D and the SNARC effect may be due simply to a generic relationship between 2D:4D and the processing of ordinal sequences represented spatially. The control assessment employed was a vowel-consonant reaction time task in which one of the letters A, B, E, F, I, J, O, P, U, V were presented to the centre of the computer screen Participants were required to identify whether the letter presented was either a vowel or consonant. Similar to the SNARC task, half of the participants responded by pressing a leftward key (Z) for vowels with the left hand and a rightward key (/) for consonants with the right hand, while the remaining participants began with the opposite key assignment (left-consonant, right-vowel). Again response key assignment was then reversed half way through the task. All other parameters of the procedure, presentation and duration were identical to that employed for the SNARC task.

6.3.4 Second to fourth digit ratio measurement

The procedure adopted in order to record and measure 2D:4D was identical to that detailed in experiment 4 with the exception that, in the current study, 2nd and 4th finger measurements taken from scanned images, were conducted by two independent raters. 2D:4D was thus computed according to averaged first and second rater 2nd and 4th finger measurements. The evaluation of successive 2nd and 4th finger measurements in experiments 1-3 suggested high intra-rater repeatability in digit measures used to calculate 2D:4D. In the current study therefore two independent raters were employed so that inter-rater repeatability of 2nd and 4th finger measurements could also be considered in the current thesis.

Intraclass correlation coefficients (r_I) suggested high inter-rater reliability for second (right hand – 0.988; left hand – 0.99) and fourth (right hand – 0.99; left hand – 0.99) fingers and 2D:4D (right hand – 0.914; left hand – 0.903) measurements. Computed TEM and rTEM measurements for the second digit were 0.57 and 0.56 (TEM) and 0.81% and 0.79% (rTEM) for the right and left hands respectively. For fourth digit measurements TEM values were calculated at 0.61 and 0.61 and rTEM values at 0.86% and 0.84% respectively for the right and left hands. Again, according to the recommendations of Weinberg et al. (2005), these values are well within an acceptable degree precision for second and fourth finger measurements.

As expected male 2D:4D ratios were found to be lower than those revealed for females for both the right (males = 0.962, SD = 0.03; females = 0.977, SD = 0.04) and left (males = 0.968, SD = 0.02; females = 0.981, SD = 0.03) hands. Despite approaching significance however sex differences in 2D:4D were not revealed to be significant for either right, $t_{(67)} = 1.936$; $p = 0.057$, or left, $t_{(67)} = 1.837$; $p = 0.071$ hand measures.

As in experiments 2 and 3, within sex median splits were applied to the data in order to facilitate the consideration of possible complex interactions between 2D:4D, task (numerical vs. control), sex and visual field. Within sex median splits according to 2D:4D for both the right and left hands were applied separately for each task data set (see table 15 for median 2D:4D values). T-test analysis confirmed that mean 2D:4D values were significantly different between each group, see Appendix 4.

Table 15

Median male and female 2D:4D values for both the right and left hands for each task data set.

	Right Hand		**Left hand**	
	Male	**Females**	**Male**	**Females**
Subitizing	0.961	0.976	0.968	0.9775
Counting	0.964	0.976	0.9695	0.977
Number Comparison	0.965	0.976	0.971	0.978
SNARC	0.963	0.977	0.971	0.978

6.3.5 Procedure

The study was approved by the School of Psychology & Sport Sciences Ethics Committee, Northumbria University. Following full informed consent, participants completed a brief biographical questionnaire followed by the computerised basic numerical tasks. To control for possible order effects tasks were completed in a random sequence. For the computerised tasks, participants were requested to lightly place their head on a chin rest positioned approximately 50cm from the centre of the monitor and focus their gaze towards the centre of the screen. Instruction for all computerised tasks emphasised both speed and accuracy. Half way through the computerised tasks, participants were given a 15min break during which they completed the Edinburgh Handedness Inventory and hand scans were obtained. Testing took approximately 1 hour 15mins. Participants were fully debriefed on completion.

6.4 Results

The means and standard deviations of reaction times (RT) to correct responses and percentage error scores were computed for each numerical task and its associated control task. Full tables of means for right and left hand 2D:4D values (for each numerical data set) and performance on each of the numerical tasks can be found in Appendix 4.

Reaction times and percentage error scores on each of the numerical tasks were explored for normality. Results of Kolmogorov-Smirnov tests used to assess normality can also be viewed in Appendix 4. For subitizing, counting and number comparison data sets findings showed that normality had been violated for a number of both reaction time and percentage error variables on either the numerical task, control task or both. Where normality had been violated, non-parametric tests were adopted in order to explore behavioural aspects of the data and simple correlations between 2D:4D and performance.

Crucially however as one of the primary points of interest was the potential complex interactions which might exist between 2D:4D, numerical task performance relative to control task and lateralization for certain numerical skills ANOVA analysis was still adopted for each task data set to evaluate any main or interaction effects of the factors digit ratio (low vs. high), sex (male vs. female), task (numerical vs. control) and, where measured, visual field of stimuli presentation (left vs. right). Separate analysis was conducted for right and left hand 2D:4D data. In line with the previous chapter, when reporting the results of each ANOVA analysis, findings relating to a potential effect of 2D:4D are reported first. Similar to experiment 2, the justification for adopting ANOVA despite violations of normaility is due to the fact that there is no non-parametric method by which complex interactions between four factors may be considered and any non-parametric one-way analysis would offer no further information to that which can be derived via simple correlational analysis. Again however it is important to consider the reported ANOVA findings in the context of the possible loss of test efficiency that may exist due to violations of normality.

With regard to the SNARC effect, in order to assist direct comparison with previous findings, the evaluation of possible 2D:4D influences on the nature of the SNARC effect was also analysed according to the procedure followed by Bull and Benson (2006), the results of this alternative analysis will not be not be reported here but can be found in Appendix 13.

6.4.1 Subitizing

6.4.1.1 Behavioural data

As normality was violated for all percentage error score measures, overall subitizing reaction times, reaction times to 2 and 3 dots stimuli and reaction times for all control variables bar responses to the colour blue, non-parametric tests were adopted in order to explore the nature of the data and simple correlations between subitizing performance and 2D:4D. Figure 11 depicts mean RT and percentage error as a function of increasing numerosity on the subitizing task and colour on the control task. Wilcoxn signed ranks analysis (Bonferroni corrected for multiple comparisons = 0.025) showed significant differences between reaction times and percentage error to the quantities 2 vs. 3 and 3 vs. 4 (see table 16). No significant differences between responses to the various colours were revealed for either reaction times or accuracy.

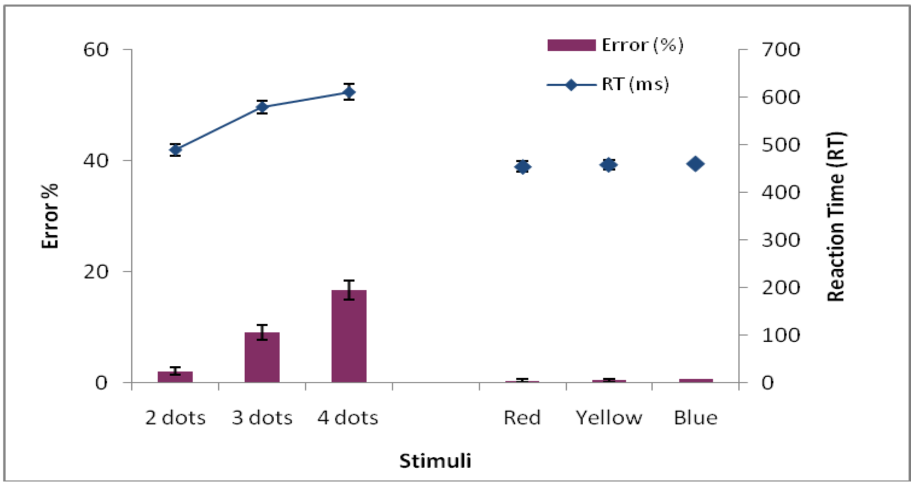


Figure 11. Mean reaction times and percentage error scores for the enumerations of quantities (2-4) in the subitizing task and colours (red, yellow and blue) in the control task, error bars indicate SEM.

Table 16

Average values and Wilcoxn signed ranks analysis of increases reaction time and percentage error as a function of ascending numerosity (n = 63), significant values indicated in bold.

	RT			Percentage error		
Quantity	Mean difference (SD)	Z	p	Mean difference (SD)	Z	p
2-3	90.07	**-6.675**	**<0.001**	6.97	**-5.198**	**<0.001**
3-4	31.37	**-3.674**	**<0.001**	7.71	**-3.843**	**<0.001**

Spearman's correlation analysis revealed no significant speed accuracy associations for performance on the subitizing control task. Significant speed accuracy associations were identified on the subitizing task in responses to stimuli presented to the RVF, ρ = 0.263; p = 0.037, and overall, ρ = 0.254; p = 0.045. Crucially however, relationships were in a positive direction, thus faster reaction times were associated with lower percentage error scores. These effects therefore are not characteristic of a speed-accuracy trade-off.

As can be seen in table 17 subitizing percentage error rates in the current experiment are more than double those revealed on the subitizing task in experiments 2 and 3. Analysis of subitizing results in the current experiment therefore considered both reaction time and accuracy (separate analysis conducted for each).

Table 17

Means and standard deviations for participant (n = 63) reaction times in ms to correct responses and percentage error for subitizing and colour recognition tasks overall and for information presented to presented to the left (LVF) and right (RVF) visual fields.

	% Error - Mean*(SD)*			**RT (ms) - Mean*(SD)***		
Task	*LVF*	*RVF*	*Overall*	*LVF*	*RVF*	*Overall*
Subitizing	9.3(*7.35*)	9.82(8.92)	9.58(7.38)	546.88(102.57)	552(105.74)	548.39(102.13)
Colours	0.57(1.01)	0.54(1.22)	0.55(0.85)	459.17(83.4)	454.17(81.14)	456.83(82.19)

6.4.1.2 Correlations

Tables 18 and 19 show the results of Spearman's correlation analysis to explore the relationship between 2D:4D and subitizing performance. Power computed via G*Power using adjusted sample sizes (see chapter 2, p. 51 for details) was again found to be low for all analyses.

As can be seen in table 18 all correlations between 2D:4D and reaction time were in a negative direction with some correlations showing small to medium effects sizes. This implies that faster reaction times were associated with high 2D:4D. With regard to percentage error scores, again small to moderate effect sizes were observed for some analyses. Analysis of percentage error scores however revealed a mixed of both positive and negative associations. No associations for analysis of either reaction times or percentage error scores in the overall sample or males and females analysed separately were found to be significant.

Table 18

Spearman correlation coefficients (ρ) for the relationship between left and right hand 2D:4D and subitizing reaction times. P values and power calculations (1-β) for each analysis are also listed.

	Males and Females n = 63		**Males n = 33**		**Females n = 30**	
Reaction times	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Subitizing Overall	ρ = -0.147 p = 0.25 1-β = 0.194	ρ = -0.144 p = 0.262 1-β = 0.188	ρ = -0.188 p = 0.294 1-β = 0.169	ρ = -0.047 p = 0.794 1-β = 0.057	ρ = -0.044 p = 0.816 1-β = 0.055	ρ = -0.102 p = 0.592 1-β = 0.079
Subitizing Left Visual Field	ρ = -0.174 p = 0.172 β = 0.255	ρ = -0.112 p = 0.38 1-β = 0.132	ρ = -0.255 p = 0.152 1-β = 0.278	ρ = -0.051 p = 0.779 1-β = 0.058	ρ = -0.08 p = 0.672 1-β = 0.068	ρ = -0.068 p = 0.723 1-β = 0.063
Subitizing Right Visual Field	ρ = -0.115 p = 0.37 1-β = 0.136	ρ = -0.16 p = 0.212 1-β = 0.222	ρ = -0.155 p = 0.389 1-β = 0.129	ρ = -0.075 p = 0.68 1-β = 0.068	ρ = -0.001 p = 0.996 1-β = 0.05	ρ = -0.12 p = 0.528 1-β = 0.091

Table 19

Spearman correlation coefficients (ρ) for the relationship between left and right hand 2D:4D and subitizing percentage error scores. P values and power calculations (1-β) for each analysis are also listed.

	Males and Females n = 63		**Males n = 33**		**Females n = 30**	
Percentage error	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Subitizing Overall	ρ = 0.088 p = 0.492 1-β = 0.1	ρ = -0.169 p = 0.186 1-β = 0.243	ρ = 0.017 p = 0.925 1-β = 0.051	ρ = -0.161 p = 0.371 1-β = 0.136	ρ = 0.116 p = 0.542 1-β = 0.088	ρ = -0.172 p = 0.364 1-β = 0.137
Subitizing Left Visual Field	ρ = 0.056 p = 0.663 1-β = 0.07	ρ = -0.233 p = 0.066 1-β = 0.42	ρ = -0.031 p = 0.864 1-β = 0.053	ρ = -0.307 p = 0.082 1-β = 0.386	ρ = 0.099 p = 0.603 1-β = 0.078	ρ = -0.157 p = 0.407 1-β = 0.122
Subitizing Right Visual Field	ρ = 0.103 p = 0.421 1-β = 0.119	ρ = -0.068 p = 0.599 1-β = 0.079	ρ = 0.072 p = 0.691 1-β = 0.066	ρ = -0.034 p = 0.85 1-β = 0.054	ρ = 0.092 p = 0.63 1-β = 0.074	ρ = -0.154 p = 0.417 1-β = 0.119

6.4.1.3 2D:4D, sex and lateralization during subitizing

Findings of the 4-Way ANOVA including the factor of digit ratio group formed on the basis of right hand 2D:4D showed no significant main effect of right hand 2D:4D and no significant interaction effects involving the factor of right hand 2D:4D for analysis of both reaction times and percentage error scores, see tables 20 (reaction times) and 21 (percentage error scores).

Table 20

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of right hand 2D:4D (all df = 1,59) for analysis of reaction times

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.007	0.934	27646.7	0.0001	0.051
2D:4D x Task interaction	2.039	0.159	5457.4	0.033	0.29
2D:4D x Sex interaction	0.002	0.967	27646.7	0.00003	0.05
2D:4D x Visual field interaction	0.529	0.47	356.818	0.009	0.11
2D:4D x Task x Sex interaction	0.58	0.449	5457.4	0.01	0.116
2D:4D x Visual field x Sex interaction	1.02	0.317	356.818	0.017	0.169
2D:4D x Task x Visual field interaction	0.77	0.783	422.374	0.001	0.059
2D:4D x Task x Sex x Visual field interaction	1.916	0.172	422.374	0.031	0.275

Table 21

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of right hand 2D:4D (all df = 1,59) for analysis of percentage error scores

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.432	0.502	58.133	0.007	0.099
2D:4D x Task interaction	0.773	0.383	54.853	0.013	0.139
2D:4D x Sex interaction	0.16	0.691	58.133	0.003	0.068
2D:4D x Visual field interaction	0.518	0.475	14.342	0.009	0.109
2D:4D x Task x Sex interaction	0.057	0.812	54.853	0.001	0.056
2D:4D x Visual field x Sex interaction	0.047	0.829	14.342	0.001	0.055
2D:4D x Task x Visual field interaction	1.338	0.252	12.394	0.022	0.207
2D:4D x Task x Sex x Visual field interaction	<0.001	0.994	12.394	<0.0001	0.05

Finding of analysis including the factor of right hand 2D:4D did reveal a significant main effect of task for analysis of both reaction time, $F_{(1,59)} = 99.999$, MSe = 5457.4, $p < 0.001$, $\eta p^2 = 0.629$, $1\text{-}\beta = 1$, and percentage error data, $F_{(1,59)} = 92.676$, MSe = 54.853, $p < 0.001$, $\eta p^2 = 0.611$, $1\text{-}\beta = 1$. Responses during the reaction time task were shown to be significantly slower and less accurate than those demonstrated for the control recognition task (see table 17). Results also revealed a significant main effect of sex on reaction times, $F_{(1,59)} = 5.959$, MSe = 27646.7, $p = 0.018$, $\eta p^2 = 0.092$, $1\text{-}\beta = 0.671$,whereby significantly slower overall response times were displayed for males (mean = 527.33, SD = 87.6) in comparison to females (mean = 475.42, SD = 74.16). No further significant main or interaction effects were revealed for analysis (including right hand 2D:4D data) of either reaction times or percentage error scores.

Analysis including left hand 2D:4D measures revealed a significant main effect of left hand 2D:4D on overall percentage error scores (see table 23), with greater accuracy being observed in high 2D:4D participants (mean = 4.08, SD = 2.32) as compared to low 2D:4D participants (mean = 6.09, SD = 4.64). As can be seen in table 23 however, analysis of percentage error data also revealed a significant left hand 2D:4D x task interaction effect. Low left hand 2D:4D (high T) participants achieved fewer errors than high left hand 2D:4D (low T) participants on the control task and higher percentage error scores in comparison to high left hand 2D:4D participants on the subitizing task (see figure 12), the main effect of left hand 2D:4D on overall percentage error scores therefore appears entirely a result of 2D:4D group differences on the subitizing task. Analysis of left hand data also revealed a significant three-way interaction between left hand 2D:4D x task and visual field on reaction times, see table 22. As can be seen in figure 13 a and b both low and high 2D:4D participant showed a right visual field advantage on the subitizing control task, this advantage however was extremely small in high 2D:4D participants. On the subitizing control task low 2D:4D participants displayed comparatively faster reaction times than high 2D:4D participants for stimuli presented to both the left and right visual fields. A similar pattern however was not demonstrated for subitizing reaction times where low 2D:4D participants showed slower reaction times than high 2D:4D participants for information presented to the right visual field. Low and high 2D:4D groups also showed opposite patterns of visual field preference on the subitizing task with a left visual field advantage observed in low 2D:4D participants and a right visual field advantage observed in high 2D:4D participants. No further significant main or interaction effect involving 2D:4D were found.

Analysis including left hand 2D:4D measures revealed a significant main effect of task for both reaction times, $F_{(1,59)} = 101.065$, MSe = 5373.69 $p < 0.001$, $\eta p^2 = 0.631$, $1-\beta = 1$, and percentage error scores, $F_{(1,59)} = 100.828$, MSe = 50.578, $p < 0.001$, $\eta p^2 = 0.631$, $1-\beta = 1$, and a significant overall main effect of sex for reaction times, $F_{(1,59)} = 5.962$, MSe = 27515.5, $p = 0.018$, $\eta p^2 = 0.092$, $1-\beta = 0.671$, all in the same direction to those described for the analysis including right hand 2D:4D groups.

Table 22

F, p, MSe, effects size (partial eta squared - ηp^2) and power ($1-\beta$) values relating to main and interaction effects for the factor of left hand 2D:4D (all df = 1,59) for analysis of reaction times (significant effects highlighted in bold).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.287	0.594	27515.5	0.005	0.082
2D:4D x Task interaction	3.467	0.068	5373.69	0.056	0.449
2D:4D x Sex interaction	0.001	0.976	27515.5	0.00002	0.05
2D:4D x Visual field interaction	0.825	0.367	359.62	0.014	0.145
2D:4D x Task x Sex interaction	0.319	0.575	5373.69	0.005	0.086
2D:4D x Visual field x Sex interaction	0.28	0.559	359.62	0.005	0.082
2D:4D x Task x Visual field interaction	**4.806**	**0.032**	403.962	0.075	0.578
2D:4D x Task x Sex x Visual field interaction	<0.001	0.995	403.962	<0.0001	0.05

Table 23

F, p, MSe, effects size (partial eta squared - ηp^2) and power ($1-\beta$) values relating to main and interaction effects for the factor of left hand 2D:4D (all df = 1,59) for analysis of percentage error scores (significant effects highlighted in bold).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	**4.676**	**0.035**	54.14	0.004	0.076
2D:4D x Task interaction	**5.442**	**0.023**	50.578	0.084	0.631
2D:4D x Sex interaction	0.231	0.632	54.14	0.004	0.076
2D:4D x Visual field interaction	0.201	0.656	14.425	0.003	0.073
2D:4D x Task x Sex interaction	0.33	0.568	50.578	0.006	0.087
2D:4D x Visual field x Sex interaction	0.046	0.831	14.425	0.001	0.055
2D:4D x Task x Visual field interaction	0.141	0.709	12.646	0.002	0.066
2D:4D x Task x Sex x Visual field interaction	<0.001	0.999	12.646	<0.0001	0.05

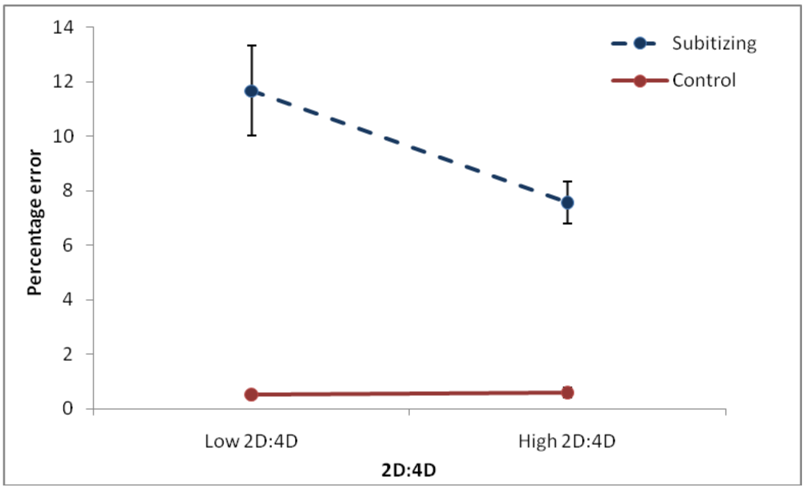
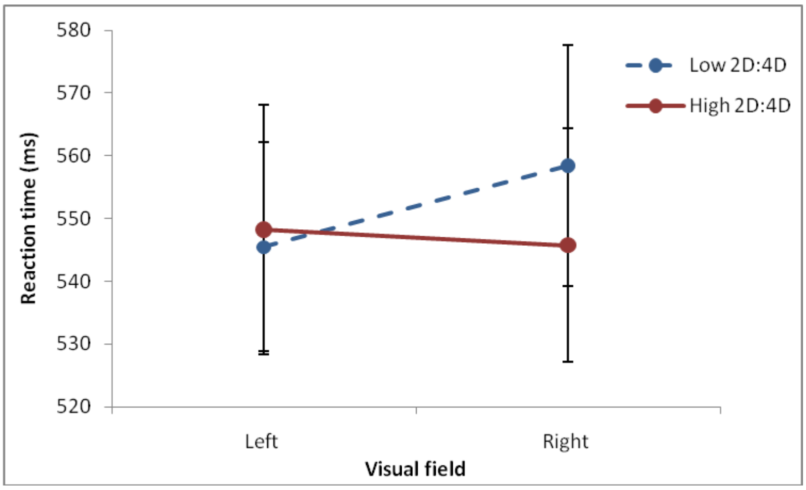


Figure 12. Average percentage error scores on the subitizing and subitizing control tasks in low and high left hand 2D:4D participants.

Post-hoc analysis of the interaction between left hand 2D:4D and task (Bonferroni corrected t-tests, $\alpha = 0.0125$) revealed significant main effects of task for both low, $t_{(30)} = 6.831$; $p < 0.001$, and high 2D:4D, $t_{(31)} = 9.163$; $p < 0.001$, participants in a similar direction to the effect revealed for overall data. While left hand 2D:4D group differences were not revealed to be significant for either the subitizing, $t_{(61)} = 2.289$; $p = 0.026$ or control task, $t_{(62)} = 0.4$; $p = 0.691$ following Bonferroni correction, such group differences were approaching significance for the subitizing task.

In order to break down the three way interaction between left hand 2D:4D, task and visual field on reaction times analysis was spilt by task and two separate two-way ANOVAs were conducted to investigate and main or interaction effects of left hand 2D:4D (low vs. high) and visual field on reaction times to the subitizing and control tasks considered separately. Separate analysis of both subitizing and control task reaction times however revealed no significant main effect of left hand 2D:4D (subitizing – $F_{(1,61)} = 0.036$, MSe = 21404.242, $p = 0.86$, $\eta p^2 = 0.001$, $1\text{-}\beta = 0.054$, control – $F_{(1,61)} = 2.017$, MSe = 13126.019, $p = 0.161$, $\eta p^2 = 0.032$, $1\text{-}\beta = 0.287$) and no significant 2D:4D x visual field interaction effects (subitizing - $F_{(1,61)} = 3.42$, MSe = 551.941, $p = 0.69$, $\eta p^2 = 0.053$, $1\text{-}\beta = 0.444$, control – $F_{(1,61)} = 1.741$, MSe = 196.276, $p = 0.192$, $\eta p^2 = 0.028$, $1\text{-}\beta = 0.255$). Analysis of control task data did reveal a significant right visual field advantage, $F_{(1,61)} = 4.083$, MSe = 13126.019, $p = 0.048$, $\eta p^2 = 0.032$, $1\text{-}\beta = 0.287$. No significant main effect of visual field was observed for subitizing reaction times, $F_{(1,61)} = 1.568$, MSe = 551.941, $p = 0.215$, $\eta p^2 = 0.025$, $1\text{-}\beta = 0.234$.

a)



b)

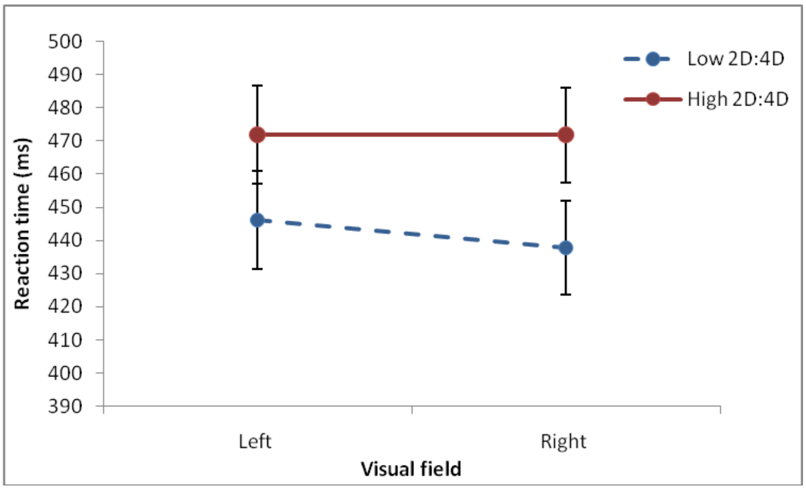


Figure 13 a & b. Mean subitizing (a) and control (b) task reaction times in low and high 2D:4D participants for responses to stimuli presented to the right and left visual fields.

6.4.1.3 Combined analysis for experiments 2-4

As can be seen in tables 18-23, similar to experiments 2 and 3, power for the analysis of subitizing data was very low. In an attempt to reconsider the potential relationship between 2D:4D and subitizing performance with a larger sample size therefore data from current study was combined with that from experiments 2 and 3.

While all three experiments considered subitizing performance the three studies possessed distinct methodologies. Firstly, the three experiments varied with regard to whether or not the stimuli were lateralized, with lateralization of stimuli adopted for experiments 2 and 4 but not for experiment 3. Composite left visual field and right visual field presentation scores therefore were not possible thus the reanalysis focused purely on overall subitizing scores (from both visual fields/ response hands combined). Secondly, the three experiments different with regard to limits placed on response time. While experiment 2 offered unlimited response time with the stimuli remaining on the screen until participant response, experiment 3 and the current study enforced both a stimuli presentation limit and a response time limit. The present experiment furthermore employed the use of a backwards mask in an attempt to limit the influence of potential retinal after images. These difference in methodology appear to have effected average subitizing reaction times for each data set with average overall subitizing response times of 891ms, 746ms, and 548, observed for experiments 2, 3 and 4 (current experiment) respectively, see tables 7 (chapter 4, p. 85), 11 (chapter 5, p. 102) and 17 (current chapter, p. 134). In order to combine the data in a meaningful manner and facilitate comparisons across the three experiments, subitizing reaction times within each dataset were converted into Z-scores. The data from each study was then combined and correlational analyses conducted in an attempt to re-investigate any associations between 2D:4D and subitizing performance. Given the low percentage error scores observed in experiment 2 and 3 (see tables 7 -chapter 4, p. 85, and 11 - chapter 5, p. 102) the analysis focused purely on reaction time data. The sample included in the analysis consisted of a total of 209 participants, 112 females and 97 males. Kolmogorov-Smirnov tests were used to investigate normality. Analysis revealed that the subitizing reaction time Z-scores were not normally distributed, see Appendix 4. Spearman's analysis therefore was adopted to investigate possible relationships between 2D:4D and subitizing scores. The results revealed no significant correlation between either left or right hand 2D:4D and subitizing performance in analysis of the entire sample, i.e. males and female combined, (right hand 2D:4D – ρ = -0.066, p = 0.343, 1-β = 0.148, left hand 2D:4D - ρ = -0.062, p = 0.374, 1-β = 0.136). Similarly no significant correlations were revealed for either male (right hand 2D:4D – ρ = -0.056, p = 0.586, 1-β = 0.082, left hand 2D:4D - ρ = 0.051, p = 0.623, 1-β = 0.076) or female data analysed separately (right hand 2D:4D - ρ = -0.067, p = 0.0482, 1-β = 0.103, left hand 2D:4D - ρ = -0.177, p = 0.062, 1-β = 0.431). A full table of means can be observed in Appendix 4. While power of the reanalysis of subitizing data was still revealed to be very low, effect sizes for each analysis were also revealed to be very low. Only one coefficient demonstrated and effect size of above 0.1, namely the negative relationship between left hand 2D:4D and subitizing reaction times observed in female data analysed separately.

6.4.2 Counting

6.4.2.1 Behavioural data

As normality was violated for all counting data variables excluding, counting reaction times to the left and right visual field, counting reaction times overall and reaction times to 7 dot stimuli, non-parametric test were adopted in order to explore the characteristics of the data. Pearson's correlations were then used to explore any potential relationships between 2D:4D and counting reaction times while Spearman's correlations were used to investigate any possible correlations between 2D:4D and counting percentage error scores. Figure 14 shows average reaction times and percentage error scores as a function of numerosity on the counting task and colour on the colour recognition task. Reaction times to 6 dot arrays were shown to be significantly faster that those to 7 dot arrays. Similarly reaction times to 7 dot arrays were significantly faster than those demonstrated in response to 8 dot arrays, see table 24. With regard to increases in percentage error with ascending numerosity, significantly more errors were revealed in response to 8 dots as compared to 7. No significant differences in percentage error however were revealed between 6 and 7 dot arrays. No significant differences in responses to the various colours on the control task were revealed for analysis of either reaction times or percentage error scores.

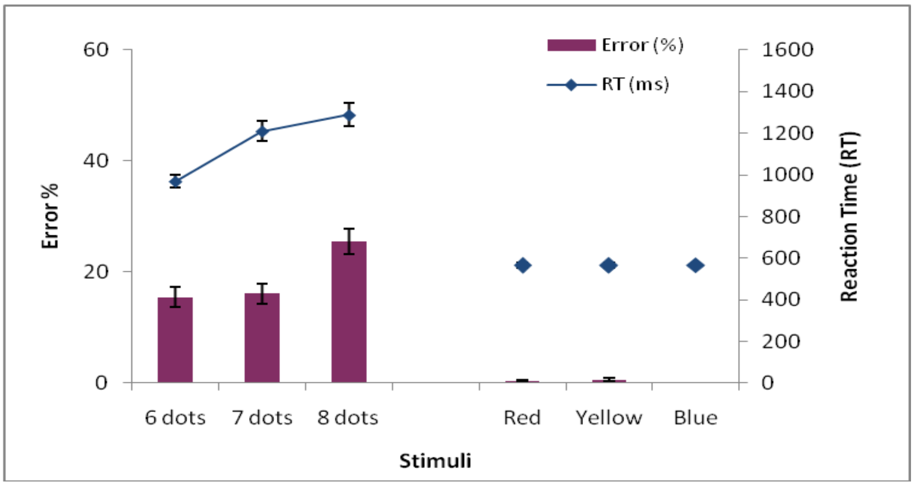


Figure 14. Mean reaction times and percentage error scores for the enumerations of quantities (6-8) in the counting task and colours (red, yellow and blue) in the control task, error bars indicate SEM.

Table 24

Average values and Wilcoxn signed ranks analysis of increases reaction time and percentage error as a function of ascending numerosity (n = 61), significant values indicated in bold.

	RT (ms)			Percentage error		
Quantity	Mean difference (SD)	Z	p	Mean difference (SD)	Z	p
6-7	204.82 (194.28)	**-6.605**	**<0.001**	0.69 (14.21)	-0.074	0.941
7-8	77.83 (171.19)	**-3.346**	**0.001**	9.44 (13.81)	**-4.538**	**<0.001**

As outlined in chapter 2 RT x set slope functions in response to the enumeration of numerosities 6-8 are typically reported at approximately 250ms – 300ms. In experiment 1, as reaction times to arrays containing 8 dots were actually lower than to those containing 7 dots it is likely that an estimation strategy as opposed to a counting strategy was adopted in the enumeration of the quantity 8. In the current study an increase in reaction time was observed with ascending numerosity from 7-8 dots. Differences in response times to the enumeration of both 6 vs. 7 and 7 vs. 8 dots however were lower than those typically reported. This difference was notably smaller for 7 vs. 8 dots. As a significant increase in percentage error with increasing numerosity was also observed in the enumeration of 7 vs. 8 dots in the absence of an increase from 6-7 dots it is possible that, similar to experiment 1, a number of participants were adopting an estimation, as opposed to counting enumeration strategy. Similar to experiment 1 therefore, in order to control for this potential confound, further analysis of the counting process included average data for the quantities 6 and 7 only.

Table 25 displays mean RTs and percentage error scores for overall performance on the counting (average scores derived from the enumeration of 6 and 7 dots) and control task. Similar to subitizing task data percentage error scores on the counting task in the current experiment were more than twice as high as those revealed in experiment 2. Analysis of counting results in the current experiment therefore also considered both reaction time and accuracy (separate analysis conducted for each).

Table 25

Means and standard deviations for participant (n = 61) reaction times in ms to correct responses and percentage error for counting and colour recognition tasks overall and for information presented to presented to the left (LVF) and right (RVF) visual fields.

	% Error - Mean*(SD)*			**RT (ms) - Mean*(SD)***		
Task	*LVF*	*RVF*	*Overall*	*LVF*	*RVF*	*Overall*
Counting	14.6 (12.67)	16.95 (13.36)	15.76 (12.29)	1085.01 (286.19)	1103.2 (209.79)	1090.24 (289.51)
Colours	0.41 (0.92)	0.41 (1)	0.41 (0.73)	567.83 (91.08)	563.43 (90.53)	566.06 (89.93)

Spearman's correlations revealed significant speed-accuracy relationships for response to the counting task overall, ρ = -0.661; p < 0.001, and for information presented to both the LVF, ρ = -0.582; p < 0.001, and RVF, ρ = -0.559; p < 0.001. All relationships were in a negative direction thus reaction times of enumeration increased with percentage error scores. No significant speed-accuracy relationships were revealed in response to the control colour recognition task.

6.4.2.2 Correlations

Tables 26 and 27 show the results of the Spearman's correlation analysis conducted in order to explore any association between 2D:4D and performance on the counting task. For analysis of reaction time data the majority of the revealed correlations were in a positive direction suggesting faster reaction times in low 2D:4D participants. The revealed coefficient values suggest small to moderate effect sizes for a number of correlations. Moderate to large effect sizes were found for relationships between left hand 2D:4D and counting performance in females. Correlations between left hand 2D:4D and counting reaction times (to both the left and right visual fields) in females were also found to be significant, see table 26. For analysis of percentage error scores a number of small to moderate effect sizes were also observed, notably however, in direct opposition to the significant correlation shown for reaction times in females, all correlations between 2D:4D measures and counting percentage error scores in females were in a positive direction. No associations between 2D:4D and percentage error scores however were found to be significant.

Table 26

Pearson's correlation coefficients (r) *for the relationship between left and right hand 2D:4D and counting reaction times. P values and power calculations (1-β) for each analysis are also listed.*

	Males and Females n = 61		**Males n = 32**		**Females n = 29**	
Reaction time	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Counting Overall	r = -0.146 p = 0.263 1-β = 0.202	r = -0.164 p = 0.207 1-β = 0.245	r = -0.113 p = 0.537 1-β = 0.094	r = 0.032 p = 0.862 1-β = 0.053	r = -0.232 p = 0.225 1-β = 0.229	**r = -0.406** **p = 0.029** **1-β = 0.608**
Counting Left Visual Field	r = -0.121 p = 0.355 1-β = 0.153	r = -0.142 p = 0.275 1-β = 0.194	r = -0.095 p = 0.606 1-β = 0.081	r = 0.076 p = 0.681 1-β = 0.07	r = -0.209 p = 0.277 1-β = 0.194	**r = -0.421** **p = 0.023** **1-β = 0.643**
Counting Right Visual Field	r = -0.15 p = 0.249 1-β = 0.212	r = -0.178 p = 0.171 1-β = 0.281	r = -0.135 p = 0.463 1-β = 0.114	r = -0.01 p = 0.955 1-β = 0.05	r = -0.229 p = 0.232 1-β = 0.224	**r = -0.394** **p = 0.035** **1-β = 0.579**

Table 27

Spearman correlation coefficients (ρ) for the relationship between left and right hand 2D:4D and counting percentage error scores. P values and power calculations (1-β) for each analysis are also listed.

	Males and Females n = 61		**Males n = 32**		**Females n = 29**	
Percentage error	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Counting Overall	ρ = 0.144 p = 0.269 1-β = 0.185	ρ = -0.07 p = 0.589 1-β = 0.08	ρ = 0.121 p = 0.51 1-β = 0.1	ρ = -0.281 p = 0.119 1-β = 0.33	ρ = 0.221 p = 0.25 1-β = 0.199	ρ = 0.223 p = 0.245 1-β = 0.202
Counting Left Visual Field	ρ = 0.168 p = 0.197 1-β = 0.237	ρ =0.075 p = 0.566 1-β = 0.085	ρ = 0.045 p = 0.807 1-β = 0.056	ρ = -0.267 p = 0.139 1-β = 0.3	ρ = 0.322 p = 0.088 1-β = 0.382	ρ = 0.21 p = 0.275 1-β = 0.184
Counting Right Visual Field	ρ = 0.121 p = 0.355 1-β = 0.144	ρ =0.033 p = 0.799 1-β = 0.057	ρ = 0.13 p = 0.48 1-β = 0.105	ρ = -0.263 p = 0.146 1-β = 0.293	ρ = 0.107 p = 0.58 1-β = 0.083	ρ = 0.188 p = 0.329 1-β = 0.156

6.4.2.3 2D:4D, sex and lateralization during counting

Using right hand 2D:4D measures in order to assess the factor of digit ratio results of the 4-Way ANOVA analysis revealed a significant interaction effect between task (counting vs. control) and right hand 2D:4D (low vs. high) on percentage error scores, see table 29. While high 2D:4D (low PT) participants showed greater accuracy than low 2D:4D (high PT) participants on the counting control task (high 2D:4D mean = 0.38, SD = 0.75; low 2D:4D mean = 0.45, SD = 0.71) the opposite pattern of results was revealed on the counting task with lower percentage error scores observed for low 2D:4D participants (high 2D:4D mean = 18.82, SD = 14.21; low 2D:4D mean = 12.6, SD = 9.12). Post hoc t-test (Bonferroni corrected for multiple comparisons, $\alpha = 0.0125$) however revealed no significant main effect of right hand 2D:4D on percentage error scores for either the counting task, $t_{(59)} = 2.025$; $p = 0.047$, or counting control task, $t_{(59)} = 0.408$; $p = 0.685$. Significant greater accuracy on the control in comparison to the counting task was revealed for both low, $t_{(29)} = 7.287$; $p < 0.001$ and high 2D:4D $t_{(30)} = 7.332$; $p < 0.001$ participants.

It is important to note however that while a significant interaction between task and right hand 2D:4D was not revealed for analysis of reaction time data, the pattern of result was in the opposite direction to that observed for percentage error scores, i.e. with faster reaction time observed for high 2D:4D participants on the counting task (high 2D:4D mean = 1053.67, SD = 276.26; low 2D:4D mean = 1128.03, SD = 302.59) and low 2D:4D participants on the control task (high 2D:4D mean =570.95, SD = 88.06; low 2D:4D mean = 561, SD = 93.06). It is possible therefore that the interaction between task and right hand 2D:4D for percentage error data may have been distorted by the influence of speed-accuracy trade-off effects. No main effects of 2D:4D and no further interaction effects involving the factor of right hand 2D:4D were revealed for analysis of either reaction time or percentage error data, see tables 28 and 29.

Table 28

F, p, MSe, effects size (partial eta squared - ηp^2) and power ($1-\beta$) values relating to main and interaction effects for the factor of right hand 2D:4D (all df = 1,57) for analysis of counting and counting control task reaction times.

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.66	0.42	104092.3	0.011	0.126
2D:4D x Task interaction	1.403	0.241	77524.75	0.024	0.214
2D:4D x Sex interaction	0.314	0.577	104092.3	0.005	0.085
2D:4D x Visual field interaction	2.05	0.158	2428.998	0.035	0.291
2D:4D x Task x Sex interaction	0.511	0.478	77524.75	0.009	0.108
2D:4D x Visual field x Sex interaction	0.071	0.791	2428.998	0.001	0.058
2D:4D x Task x Visual field interaction	1.539	0.22	2520.95	0.026	0.23
2D:4D x Task x Sex x Visual field interaction	0.047	0.829	2520.95	0.001	0.055

Table 29

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of right hand 2D:4D (all df = 1,57) for analysis of counting and counting control task percentage error scores (significant effect indicated in bold).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	3.783	0.057	150.818	0.062	0.481
2D:4D x Task interaction	**4.175**	**0.046**	143.491	0.068	0.52
2D:4D x Sex interaction	0.067	0.796	150.818	0.001	0.057
2D:4D x Visual field interaction	0.248	0.621	19.705	0.004	0.078
2D:4D x Task x Sex interaction	0.054	0.818	143.491	0.001	0.056
2D:4D x Visual field x Sex interaction	1.198	0.278	19.705	0.021	0.19
2D:4D x Task x Visual field interaction	0.654	0.422	20.634	0.011	0.125
2D:4D x Task x Sex x Visual field interaction	0.245	0.622	20.634	0.004	0.078

4-way ANOVA analysis including right hand 2D:4D measures revealed an overall main effect of task on both reaction times, $F_{(1,57)} = 221.554$, MSe = 77524.75, $p < 0.001$, $\eta p^2 = 0.795$, 1-β = 1, and percentage error scores, $F_{(1,57)} = 98.474$, MSe = 143.491. $p < 0.001$, $\eta p^2 = 0.633$, 1-β = 1, with responses to the counting task shown to be significantly slower and less accurate than those to the control colour recognition task, see table 25. Analysis of percentage error scores also revealed a significant main effect of visual field, $F_{(1,57)} = 4.319$, MSe = 19.705, $p = 0.042$, $\eta p^2 = 0.07$, 1-β = 0.533, and a significant interaction between task (counting vs. control) and visual field, $F_{(1,57)} = 4.103$, MSe = 20.634, $p = 0.047$, $\eta p^2 = 0.067$, 1-β = 0.513. On average, responses to stimuli presented to the left visual field were significantly more accurate in comparison those displayed for stimuli presented to the right visual field, as can be seen in table 25 however, percentage error scores on the control task were similar for stimuli presented to both the left and right visual field. Overall average differences therefore appear to be a consequence of visual field difference on the counting task.

While a significant task x visual field interaction was not observed for reaction time, $F_{(1,57)} = 3.212$, MSe = 2520.95, $p = 0.078$, $\eta p^2 = 0.053$, 1-β = 0.422, average reaction times for information presented to the left visual field in comparison to the right visual field were lower on the counting task and high on the control task, see table 25. The significant interaction between task and visual field revealed for analysis of percentage error data therefore cannot be easily explained with reference to speed-accuracy trade-off effects. Post-hoc analysis of the revealed interaction between task and visual field (Bonferroni corrected t-tests) on percentage error scores demonstrated a significant effect of task for stimuli presented to both visual fields (LVF – $t_{(60)} = 8.75$; p

< 0.001, RVF – $t_{(60)} = 9.763$; $p < 0.001$), in line with the overall effects of task described above. Visual field differences for both the counting ($t_{(60)} = 2.095$; $p = 0.04$) and control task ($t_{(60)} = 0.004$; $p = 0.996$) however were not found to be significant following Bonferroni correction ($\alpha = 0.0125$).

Findings of the 4-Way ANOVA using left hand 2D:4D measures revealed a significant task x left hand 2D:4D interaction on reaction times, see table 30. In contrast to the analysis including right hand 2D:4D measures the same interaction was not found to be significant for analysis of percentage error scores. As can be seen in figure 15, While high left hand 2D:4D (low PT) participants showed faster reaction times in comparison to low left hand 2D:4D (high PT) participants on the counting task, the opposite pattern of results was revealed for the control task. While not significant, patterns of percentage error data across high (counting mean = 15.62, SD = 12.31, control mean = 0.48, SD = 0.81) and low (counting mean = 15.91, SD = 12.47, control mean = 0.35, SD = 0.64) 2D:4D participants were in a similar direction thus it is unlikely that this interaction can be accounted for with reference to simple speed accuracy trade-off effects. Post-hoc analysis (Bonferroni corrected for multiple comparisons, $\alpha = 0.125$) revealed a significant main effect of task for both high 2D:4D, $t_{(30)} = -11.383$; $p < 0.001$, and low 2D:4D, $t_{(29)} = -10.188$; $p < 0.001$, participant data analysed independently, with both group demonstrating significantly faster reaction times on the control task. No significant main effect of left hand digit ratio group however was revealed for either the counting, $t_{(59)} = 1.36$; $p = 0.179$ or control, $t_{(59)} = 1.883$; $p = 0.065$, task.

Table 30

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of left hand 2D:4D (all df = 1,57) for analysis of counting and counting control task reaction times (significant effect indicated in bold).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.695	0.408	101114.8	0.012	0.13
2D:4D x Task interaction	**4.79**	**0.033**	73030.731	0.078	0.576
2D:4D x Sex interaction	2.055	0.157	101114.8	0.035	0.291
2D:4D x Visual field interaction	1.444	0.235	2403.775	0.025	0.291
2D:4D x Task x Sex interaction	0.852	0.36	73030.731	0.015	0.148
2D:4D x Visual field x Sex interaction	1.203	0.277	2403.775	0.021	0.19
2D:4D x Task x Visual field interaction	1.361	0.248	2394.009	0.023	0.209
2D:4D x Task x Sex x Visual field interaction	3.147	0.081	2394.009	0.052	0.415

Table 31

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of left hand 2D:4D (all df = 1,57) for analysis of counting and counting control task percentage error scores.

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.001	0.975	158.967	0.00002	0.05
2D:4D x Task interaction	0.011	0.916	151.627	0.002	0.051
2D:4D x Sex interaction	0.687	0.411	158.967	0.012	0.129
2D:4D x Visual field interaction	0.209	0.649	19.324	0.004	0.073
2D:4D x Task x Sex interaction	0.885	0.351	151.627	0.015	0.152
2D:4D x Visual field x Sex interaction	2.4	0.127	19.324	0.04	0.331
2D:4D x Task x Visual field interaction	0.045	0.833	20.134	0.001	0.055
2D:4D x Task x Sex x Visual field interaction	2.279	0.137	20.134	0.038	0.317

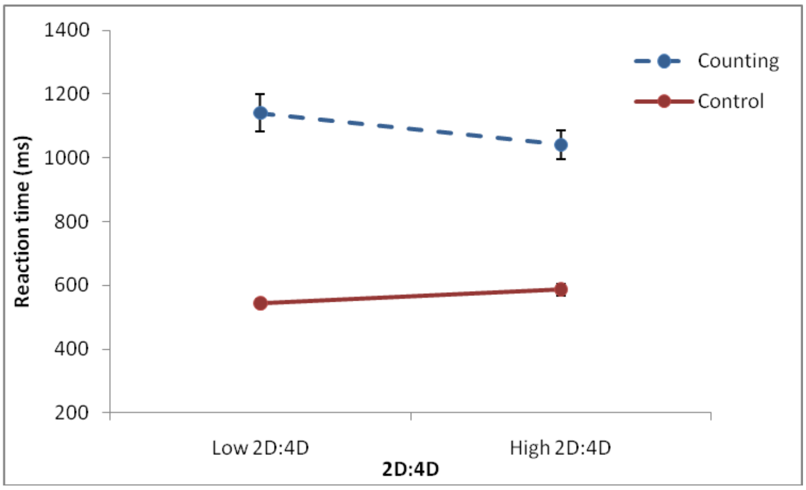


Figure 16. Average percentage error scores on the counting and counting control tasks in low and high left hand 2D:4D participants.

Findings of the 4-Way ANOVA using left hand 2D:4D measures also revealed a main effect of task on both reaction times, $F_{(1,57)} = 235.798$, MSe = 73030.731, $p < 0.001$, $\eta p^2 = 0.805$, 1-β = 1, and percentage error scores, $F_{(1,57)} = 93.649$, MSe = 151.627, $p < 0.001$, $\eta p^2 = 0.622$, 1-β = 1, as well as a significant main effect of visual field on percentage error scores, $F_{(1,57)} = 4.146$, MSe = 19.324, $p = 0.046$, $\eta p^2 = 0.068$, 1-β = 0.517. The significant task x visual field interaction revealed in analysis including the factor of right hand 2D:4D was also approaching significance in the analysis including the factor of left hand 2D:4D data, $F_{(1,57)} = 3.993$, MSe = 20.134, $p = 0.05$, $\eta p^2 = 0.065$, 1-β = 0.502. All such effects were identical in nature to those described for analysis including right hand 2D:4D measures.

6.4.2.4 Combined analysis for experiments 2 and 4

As can been seen in tables 26-31 low power was again revealed for both correlation and ANOVA analysis of the counting data set. As the process of counting was assessed in both experiments 2 and 4 a similar procedure to that employed in section 6.2.1.4 above for subitizing data was thus applied to counting. Data from experiment 2 and the current study was therefore combined and correlation analysis conducted in order to re-explore the association between left and right hand 2D:4D and counting performance using larger sample sizes. Combination of data across the two experiments resulted in a sample size of 137 participants, 75 males and 62 females. Although the presentation of counting stimuli across the two experiments was similar (both experiments lateralized stimuli and offered unlimited stimuli presentation and response times), analysis of overall counting reaction times suggested that mean reaction times were different across the two studies, $t(135) = 9.103$, $p < 0.001$, see tables 7 (chapter 4, p. 85) and 25 (current chapter, p. 143) for means. A similar procedure to that adopted for the re-analysis of subitizing data therefore was employed whereby raw counting reaction times within each dataset were converted into Z-scores and the re-analysis conducted on the basis of this standardised data. As stimuli from both experiments was lateralized analysis was conducted in order to explore relationships between 2D:4D and reaction times overall and for information presented to the left and right visual field separately. Given the low percentage error scores in experiment 2 (see table 7, chapter 4, p. 85) the analysis focused exclusively on reaction times. Kolmogorov-Smirnov analysis revealed that all data was normally distributed (see Appendix 4) thus Pearson's correlation analysis was conducted in order to explore associations. As can be seen from table 32 all correlations with the exception of those involving left hand 2D:4D in males were in a negative direction (i.e. high 2D:4D was associated with faster reaction times). Correlations in females and in males between left hand 2D:4D and left visual field counting reaction times showed a small to moderate effect size. Interestingly however the coefficient sizes observed in females were reduced relative to those observed when considering data from the two experiments separately. Overall no significant correlations were revealed between either right or left hand 2D:4D and counting performance. Full tables of means for the variables relating to this analysis can be viewed in Appendix 4.

Table 32

Pearson's correlation coefficients (r) *for the relationship between left and right hand 2D:4D and Z scores for counting reaction times for analysis of data from experiment 2 and the current study combined. P values and power calculations (1-β) for each analysis are also listed.*

	Males and Females n = 137		**Males n = 62**		**Females n = 75**	
Reaction time	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Counting Overall	r = -0.074 p = 0.388 1-β = 0.138	r = -0.034 p = 0.698 1-β = 0.068	r = -0.054 p = 0.677 1-β = 0.07	r = 0.07 p = 0.586 1-β = 0.08	r = -0.168 p = 0.149 1-β = 0.304	r = -0.175 p = 0.134 1-β = 0.326
Counting Left Visual Field	r = -0.057 p = 0.506 1-β = 0.101	r = -0.008 p = 0.929 1-β = 0.051	r = -0.029 p = 0.826 1-β = 0.056	r = 0.106 p = 0.414 1-β = 0.13	r = -0.157 p = 0.179 1-β = 0.271	r = -0.156 p = 0.182 1-β = 0.268
Counting Right Visual Field	r = -0.082 p = 0.34 1-β = 0.159	r = -0.057 p = 0.507 1-β = 0.101	r = -0.077 p = 0.554 1-β = 0.091	r = 0.039 p = 0.761 1-β = 0.06	r = -0.172 p = 0.141 1-β = 0.316	r = -0.197 p = 0.09 1-β = 0.399

6.4.3 Number comparison

6.4.3.1 Behavioural data

Normality in the number comparison data set was violated for all variables except; reaction times to 0.57 and 0.67 comparison ratios on the control task, percentage error scores for the 0.67 and 0.8 comparison ratios on the number comparison task and overall number comparison percentage scores for stimuli presented to the left visual field, right visual field and overall (data from both visual fields combined). Non parametric test were therefore adopted to explore the characteristics of the data and associations between 2D:4D and number comparison reaction times. Pearson's correlations were used to explore associations between 2D:4D and number comparison percentage error scores. Table 33 shows means and standard deviations for reaction times to correct responses and percentage error scores on the number comparison and number comparison control task. Significant speed-accuracy relationships were evident in response to the number comparison task overall, ρ = -0.446; p < 0.001, and for information presented to both the LVF, ρ = -0.367; p = 0.002, and RVF, ρ = -0.507; p < 0.001. Significant speed-accuracy associations were also revealed for performance on the number comparison control task overall, ρ = -0.487; p < 0.001, and for information presented to both the LVF, ρ = -0.427; p < 0.001, and RVF, ρ = -0.325; p = 0.008. All relationships for analysis of both tasks were in a negative direction thus increased reaction times were associated with increased percentage error scores.

Table 33

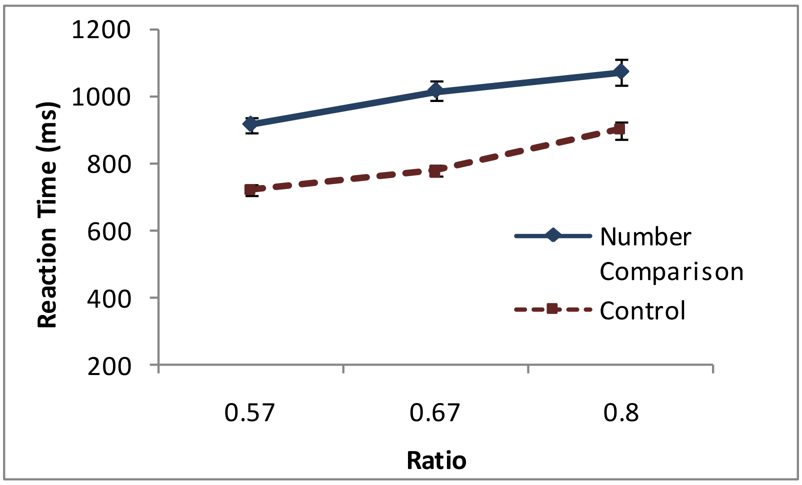
Means and standard deviations for participant (n = 66) reaction times in ms to correct responses and percentage error for the number comparison and number comparison control tasks overall and for information presented to the left and right visual fields independently.

		Visual field		
Task		**Left**	**Right**	**Overall**
Number Comparison	Mean RT in ms (SD)	978.11 (234.29)	985.3 (237.47)	982.05 (234.81)
	Mean % Error (SD)	20.23 (7.73)	20.52 (7.67)	20.23 (7.15)
Control	Mean RT in ms (SD)	787.95 (145.28)	777.52 (138.52)	783.11 (143.35)
	Mean % Error (SD)	3.51 (4.13)	4.67 (4.77)	4.09 (3.75)

As can be seen in figure 16 a and b, in line with expected distance effects both reaction times and percentage error scores increased as the ratio between the two quantities/square sizes under comparison decreased. Wilcoxn signed ranks analysis showed significant differences between: reaction times to responses for 0.57 ratio comparisons vs. 0.67 ratio comparisons, Z = -6.596; p < 0.001; reaction times to responses for 0.67 ratio comparisons vs. 0.8 ratio comparisons, Z = -3.089; p = 0.002 on the number comparison task. Analyses also showed significant differences in number comparison percentage error scores to responses for 0.57 ratio comparisons vs. 0.67 ratio comparisons, Z = -6.714, p < 0.001; and percentage error scores to responses for 0.67 ratio comparisons vs. 0.8 ratio comparisons, Z = -7.056 ; p < 0.001, on the number comparison task.

With reference to the control task, significant differences were evident between: reaction times to responses for 0.57 ratio comparisons vs. 0.67 ratio comparisons, Z = -6.449; p < 0.001; reaction times to responses for 0.67 ratio comparisons vs. 0.8 ratio comparisons, Z = -6.785; p < 0.001; and percentage error scores to responses for 0.67 ratio comparisons vs. 0.8 ratio comparisons, Z = -5.999; p < 0.001. No significant differences however were found between responses for 0.57 ratio comparisons vs. 0.67 ratio comparisons for percentage error scores on the control task, Z = -1.038; p = 0.299.

a)



b)

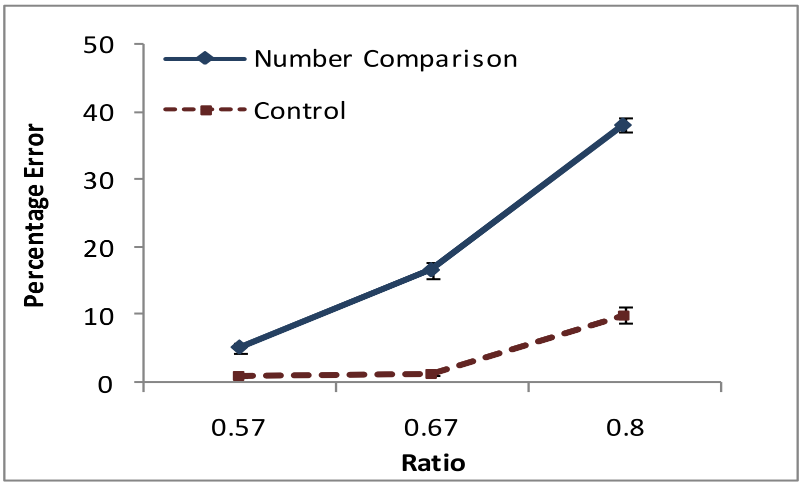


Figure 16 a and b. Mean reaction times (a) for correct responses (ms) and percentage error scores (b) for performance on the number comparison and control tasks over each ratio level of difference between number and size comparisons, including error bars indicating SEM.

6.4.3.2 Correlations

Tables 34 and 35 display results of correlation analyses to investigate any simple associations between 2D:4D and number comparison task performance. All correlations excluding those revealed for left hand 2D:4D in males were in a negative direction thus higher 2D:4D was associated with faster reaction times. A number of small to medium effects were observed. For analyses of percentage error scores the majority of effect sizes were very low. Where small to medium effect were observed (i.e. between measure of number comparison performance and right hand 2D:4D in males and between left and right hand 2D:4D and right visual field reaction times in females) effect were in a positive direction in males and a negative direction in females. Power values for each analysis were again shown to be very low. None significant association were found for overall data or male and female data analysed independently.

Table 34

Spearman correlation coefficients (ρ) for the relationship between left and right hand 2D:4D and number comparison reaction times. P values and power calculations (1-β) for each analysis are also listed.

	Males and Females n = 66		**Males n = 33**		**Females n = 33**	
Reaction time	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
No. Comp Overall	ρ = -0.203 p = 0.102 1-β = 0.35	ρ = -0.079 p = 0.527 1-β = 0.093	ρ = -0.15 p = 0.406 1-β = 0.127	ρ = 0.126 p = 0.485 1-β = 0.103	ρ = -0.188 p = 0.295 1-β = 0.174	ρ = -0.154 p = 0.392 1-β = 0.131
No. Comp Left Visual Field	ρ = -0.141 p = 0.258 1-β = 0.192	ρ = -0.043 p = 0.73 1-β = 0.062	ρ = -0.062 p = 0.731 1-β = 0.063	ρ = 0.156 p = 0.385 1-β = 0.133	ρ = -0.085 p = 0.639 1-β = 0.074	ρ = -0.124 p = 0.492 1-β = 0.102
No. Comp Right Visual Field	ρ = -0.221 p = 0.075 1-β = 0.406	ρ = -0.112 p = 0.37 1-β = 0.138	ρ = -0.121 p = 0.503 1-β = 0.1	ρ = 0.109 p = 0.546 1-β = 0.09	ρ = -0.224 p = 0.21 1-β = 0.229	ρ = -0.254 p = 0.154 1-β = 0.284

Table 35

Pearson's correlation coefficients (r) *for the relationship between left and right hand 2D:4D and number comparison percentage error scores. P values and power calculations (1-*β*) for each analysis are also listed.*

	Males and Females n = 66		**Males n = 33**		**Females n = 33**	
Percentage error	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
No. Comp Overall	r = 0.051 p = 0.683 1-β = 0.069	r = 0.014 p = 0.911 1-β = 0.051	r = 0.106 p = 0.557 1-β = 0.09	r = -0.05 p = 0.781 1-β = 0.059	r = -0.078 p = 0.668 1-β = 0.071	r = -0.078 p = 0.668 1-β = 0.071
No. Comp Left Visual Field	r = 0.088 p = 0.483 1-β = 0.108	r = 0.07 p = 0.575 1-β = 0.086	r = 0.109 p = 0.545 1-β = 0.092	r = 0.02 p = 0.912 1-β = 0.051	r = 0.012 p = 0.947 1-β = 0.05	r = 0.046 p = 0.8 1-β = 0.057
No. Comp Right Visual Field	r = 0.010 p = 0.936 1-β = 0.051	r = -0.036 p = 0.777 1-β = 0.059	r = 0.118 p = 0.514 1-β = 0.1	r = -0.062 p = 0.733 1-β = 0.063	r = -0.163 p = 0.366 1-β = 0.148	r = -0.112 p = 0.536 1-β = 0.095

6.4.3.3 2D:4D, sex and lateralization during number comparison

The results of the 4-way ANOVA utilising right hand 2D:4D measures showed a significant task x digit ratio group interaction effect on reaction times, see table 36. As can be seen in figure 17, both low and high 2D:4D participants showed faster reaction times on the control as compared to number comparison task. While high 2D:4D (low PT) participants however demonstrated comparatively faster reaction times on the number comparison task as compared low 2D:4D participants (high PT) the opposite pattern of results was displayed for the control task (faster reaction times observed in low 2D:4D participants). No significant task x right hand 2D:4D interaction was observed for analysis of percentage error data, crucially however while patterns of performance on the control task according to percentage error data was similar to those shown when considering reaction times, i.e. fewer errors observed for low 2D:4D participants (low 2D:4D mean = 3.95, SD = 3.74, high 2D:4D mean = 4.21, SD = 3.81), patterns of percentage error on the number comparison task did not mirror those revealed for reaction times. During the number comparison task high 2D:4D (mean = 21.02, SD = 7.25) participants showed reduced accuracy in comparison to low 2D:4D participants (mean = 19.33, SD = 7.04) implying potential speed accuracy effects on number comparison task performance which may have influenced the revealed interaction between task and right hand 2D:4D on reaction times.

Post-hoc t-tests (Bonferroni corrected for multiple comparison, $\alpha = 0.0125$) of the significant interaction between task and right hand 2D:4D on reaction times revealed a significant main effect of task for both low, $t_{(30)} = 8.269$; $p < 0.001$, and high 2D:4D, participants, $t_{(34)} = 8.167$; $p < 0.001$, with significantly faster reaction times observed on the control as compared to the number comparison task. No significant 2D:4D group differences however were revealed for reaction times on either the number comparison, $t_{(64)} = 1.47$; $p = 0.146$, or control task, $t_{(64)} = 0.414$; $p = 0.68$.

As can be seen in tables 36 and 37, no futher main effect or interaction effect involving right hand 2D:4D were identified in analysis of either reaction times or percentage error scores.

Table 36

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of right hand 2D:4D (all df = 1,57) for analysis of number comparison and number comparison control task reaction times (significant effect indicated in bold).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.468	0.497	119537.19	0.007	0.103
2D:4D x Task interaction	**7.085**	**0.01**	20618.702	0.103	0.746
2D:4D x Sex interaction	0.079	0.779	119537.19	0.001	0.059
2D:4D x Visual field interaction	0.843	0.362	1445.129	0.013	0.148
2D:4D x Task x Sex interaction	0.216	0.644	20618.702	0.003	0.074
2D:4D x Visual field x Sex interaction	0.987	0.324	1445.129	0.016	0.165
2D:4D x Task x Visual field interaction	1.545	0.219	918.996	0.024	0.232
2D:4D x Task x Sex x Visual field interaction	0.001	0.975	918.996	0.00002	0.05

Table 37

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of right hand 2D:4D (all df = 1,57) for analysis of number comparison and number comparison control task percentage error scores.

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.862	0.357	87.448	0.014	0.15
2D:4D x Task interaction	1.074	0.304	39.836	0.017	0.175
2D:4D x Sex interaction	0.287	0.594	87.448	0.005	0.082
2D:4D x Visual field interaction	0.487	0.488	12.601	0.008	0.106
2D:4D x Task x Sex interaction	0.341	0.561	39.836	0.005	0.089
2D:4D x Visual field x Sex interaction	0.18	0.673	12.601	0.003	0.07
2D:4D x Task x Visual field interaction	0.287	0.594	19.398	0.005	0.082
2D:4D x Task x Sex x Visual field interaction	0.072	0.789	19.398	0.001	0.058

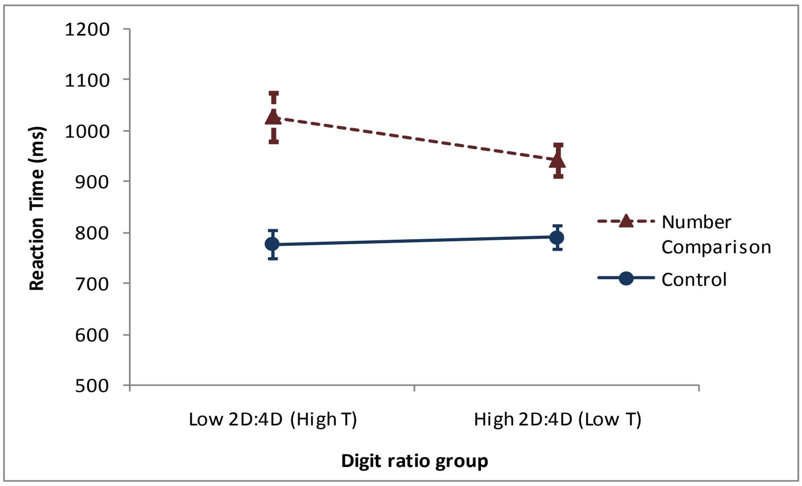


Figure 17. Mean number comparison and control task reaction times for low (high testosterone) and high (low testosterone) right hand 2D:4D participants, including error bars indicating SEM.

Analysis of reaction time data using right hand 2D:4D groups revealed a significant overall main effect of task on both response times, $F_{(1,62)} = 129.469$, MSe = 20618.702 $p < 0.001$, $\eta p^2 = 0.676$, $1\text{-}\beta = 1$, and percentage error scores, $F_{(1,62)} = 435.56$, MSe = 39.836, $p < 0.001$, $\eta p^2 = 0.875$, $1\text{-}\beta = 1$, with significantly faster reaction times and higher accuracy revealed for the control task as compared to the number comparison task, see table 33. For analysis of reaction times however, a significant main effect of sex was also observed, $F_{(1,62)} = 6.241$, MSe = 119537.19, $p = 0.015$, $\eta p^2 = 0.091$, $1\text{-}\beta = 0.691$, with males demonstrating significantly slower overall reaction times (mean = 936.16, SD = 199.45) than females (mean = 828.99, SD = 141.24). Although not significant, $F_{(1,62)} = 1.851$, MSe = 87.448, $p = 0.179$, $\eta p^2 = 0.029$, $1\text{-}\beta = 0.268$, males however did displayed lower overall percentage error scores (mean = 11.3, SD = 4.91) in comparison to females (mean = 13.01, SD = 4.4). The potential influence of speed accuracy effects on sex differences in reaction times therefore cannot be dismissed. Analysis of reaction times using right hand 2D:4D measures also revealed a significant task x visual field interaction, $F_{(1,62)} = 5.915$, MSe = 918.996, $p = 0.018$, $\eta p^2 = 0.087$, $1\text{-}\beta = 0.668$.

While a left visual field (right hemisphere) advantage was observed in reaction times to the number comparison task, a right visual field advantage was revealed in reaction times to the control task (see table 33). Post-hoc t-tests (Bonferroni corrected for multiple comparisons, $\alpha = 0.0125$) of the significant interaction between task and visual field on reaction times showed an overall main effect of task on reaction times for information presented to both the left, $t_{(65)} = 10.3$; $p < 0.001$, and right visual fields, $t_{(65)} = 10.78$; $p < 0.001$, with significantly faster response times evident for the control task, see table 21. Visual field effects for both the number comparison, $t_{(65)} = 1.002$; $p = 0.32$, and control task, $t_{(65)} = 2.267$; $p = 0.027$, however were not revealed to be significant following Bonferroni correction ($\alpha = 0.0125$). While a significant task x visual field interaction effect was not observed for analysis of percentage error data, $F_{(1,62)} = 0.612$, MSe = 19.398, $p = 0.437$, $\eta p^2 = 0.01$, $1\text{-}\beta = 0.12$, patterns of percentage error on the number comparison task were similar to those revealed for reaction time, i.e. LVF advantage. Notably however while a significant right visual field advantage was observed for reaction times to the control task percentage errors scores showed a left visual field advantage, see table 33. Again therefore, it is possible that potential speed-accuracy trade-off effects may have influenced the observed interaction between task and visual field on reaction time.

No significant main effect of left hand 2D:4D or significant interaction effect involving left hand 2D:4D were identified for analysis of either reaction time or percentage error data, see tables 38 and 39. 4-way ANOVA analysis including left hand 2D:4D measures revealed a similar main effect of task on both reaction times, $F_{(1,62)}$ = 117.479, MSe = 22332.521, p < 0.001, ηp^2 = 0.655, 1-β = 1, and percentage error, $F_{(1,62)}$ = 431.027, MSe = 40.536, p < 0.001, ηp^2 = 0.874, 1-β = 1, a similar main effect of sex on reaction times, $F_{(1,62)}$ = 6.219, MSe = 119773.75 p = 0.015, ηp^2 = 0.091, 1-β = 0.69, and a similar task x visual field interaction on reaction time, $F_{(1,61)}$ = 5.977, MSe = 891.449, p = 0.017, ηp^2 = 0.088, 1-β = 0.672. The nature of all such effects were identical to those described above for analysis including right hand 2D:4D data.

Table 38

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of left hand 2D:4D (all df = 1,57) for analysis of number comparison and number comparison control task reaction times (significant effect indicated in bold).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.311	0.579	119773.75	0.005	0.085
2D:4D x Task interaction	0.802	0.374	22332.521	0.013	0.143
2D:4D x Sex interaction	0.114	0.736	119773.75	0.002	0.063
2D:4D x Visual field interaction	0.002	0.969	1487.901	0.00002	0.05
2D:4D x Task x Sex interaction	1.189	0.28	22332.521	0.019	0.189
2D:4D x Visual field x Sex interaction	<0.001	0.983	1487.901	0.0004	0.05
2D:4D x Task x Visual field interaction	2.35	0.13	891.449	0.037	0.326
2D:4D x Task x Sex x Visual field interaction	1.16	0.286	891.449	0.018	0.185

Table 39

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of left hand 2D:4D (all df = 1,57) for analysis of number comparison and number comparison control task percentage error scores.

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	<0.001	0.996	88.333	<0.0001	0.272
2D:4D x Task interaction	0.128	0.722	40.536	0.002	0.064
2D:4D x Sex interaction	0.527	0.473	88.333	0.008	0.11
2D:4D x Visual field interaction	0.611	0.438	12.598	0.01	0.12
2D:4D x Task x Sex interaction	0.196	0.659	40.536	0.003	0.072
2D:4D x Visual field x Sex interaction	0.073	0.788	12.598	0.001	0.058
2D:4D x Task x Visual field interaction	0.263	0.61	19.425	0.004	0.08
2D:4D x Task x Sex x Visual field interaction	0.001	0.921	19.425	0.0002	0.051

6.4.4 SNARC

6.4.4.1 Behavioural data

Average reaction times to correct responses and percentage error scores were calculated. No significant speed-accuracy associations were revealed for either SNARC task, ρ = 0.161, p = 0.194, or vowel-consonant task performance, ρ = 0.098, p = 0.43. Notably however mean percentage error scores were low with a number of participants performing at or close to ceiling (SNARC mean = 4.388, SD = 3.7; vowel-consonant mean = 3.63, SD = 3.28). Only reaction times therefore were considered in the analysis.

Following the same procedure as Bull and Benson (2006) the SNARC effect was evaluated using reaction time difference scores, calculated as left hand RT minus right hand RT for each digit/letter and separately for each participant. On the SNARC task it was predicted that response times would reflect a spatially organised representation of numerical magnitude with reaction times to low digits demonstrated to be faster with the left hand and responses to higher digits revealed to be faster with the right hand. Similarly on the vowel-consonant control task it was expected that response latencies would reflect a spatially organised ordinal representation of letters of the alphabet, thus letters closer to the beginning of the alphabet would be responded to faster with the left hand and letters towards the end of the alphabet would be responded to faster with the right hand. As reaction time difference scores were calculated as right hand reaction time minus left hand reaction time, it was in turn anticipated that responses to low digits and letters at the beginning of the alphabet would show a positive reaction time difference score while responses to higher digits and letters at the end of the alphabet would show a negative reaction time differences.

For each participant's data the calculated reaction time difference scores were used in a repeated measures regression analysis with digit magnitude as the predictor variable and RT difference as the criterion variable in order to compute a regression equation. The resulting regression weight for each participant was then recorded. Analysis of all data using a one sample t-test confirmed the presence of an overall SNARC effect with regression weights revealed to differ significantly from 0, $t_{(66)}$ = 4.389; p < 0.001. In analysis of all data on the vowel-consonant task however regression weights did not differ significantly from 0, $t_{(66)}$ = 1.501; p = 0.138. The task therefore appears not to have effectively tapped an ordinal representation of letters of the alphabet.

6.4.4.2 Correlations

As calculated regression weights for both the SNARC and vowel-consonant task were found to be normally distributed, see appendix 4, Pearson's correlation analyses were use to investigate any possible relationships between 2D:4D and the magnitude of the SNARC effect. While results revealed small to moderate effect correlation coefficients for analysis of male data in a positive direction (i.e. with lower 2D:4D associated with smaller regression weights) no significant correlations were revealed between either right or left hand 2D:4D and SNARC regression weight in the overall sample (right hand – $r = 0.029$, $p = 0.817$, $1\text{-}\beta = 0.056$, left hand – $r = 0.154$, $p = 0.215$, $1\text{-}\beta = 0.239$) or male (right hand – $r = 0.118$, $p = 0.5$, $1\text{-}\beta = 0.103$, left hand – $r = 0.276$, $p = 0.109$, $1\text{-}\beta = 0.367$) and female data analysed separately (right hand – $r = -0.051$, $p = 0.782$, $1\text{-}\beta = 0.059$, left hand – $r = 0.068$, $p = 0.711$, $1\text{-}\beta = 0.066$).

6.4.4.3 2D:4D, sex and SNARC

A 3-way ANOVA was conducted in order to explore any main or interaction effects of the factors; task (SNARC vs. vowel-consonant control task), 2D:4D (low vs. high) and sex (male vs. female) on the calculated regression weights. For comparison with previous literature however the evaluation of possible 2D:4D influences on the nature of the SNARC effect was also analysed according to the procedure followed by Bull and Benson (2006), the results of this alternative analysis can be found in Appendix 13.

As can be seen in table 40, no significant main effect of either right or left hand 2D:4D or significant interaction effects involving right or left hand 2D:4D were identified. A significant main effect of task was revealed in the analysis including left hand 2D:4D, $F_{(1,63)} = 4.969$, $MSe = 0.105$, $p = 0.029$, $\eta p^2 = 0.076$, $1\text{-}\beta = 0.609$, and right hand 2D:4D, $F_{(1,63)} = 5.164$, $MSe = 0.105$, $p = 0.026$, $\eta p^2 = 0.073$, $1\text{-}\beta = 0.593$. Average regression weights were higher for the SNARC task (SNARC mean = 0.19, SD = 0.35, vowel-consonant mean = 0.06, SD = 0.32) suggesting a stronger ordinal representation of numbers on the SNARC task that letters of the alphabet on the vowel-consonant task. No further significant main or interaction effect were identified.

Table 40

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of 2D:4D (separate analysis for left and right hand 2D:4D) for analysis of SNARC and vowel-consonant task regression weights (all df = 1,63).

Analysis	**Effect**	**F**	**p**	**MSe**	ηp^2	**1-β**
Right Hand 2D:4D	Main effect 2D:4D	0.961	0.331	0.124	0.015	0.162
	2D:4D x Task interaction	2.991	0.089	0.105	0.045	0.399
	2D:4D x Sex interaction	0.123	0.727	0.124	0.002	0.064
	2D:4D x Task x Sex interaction	0.011	0.918	0.105	0.0002	0.051
Left Hand 2D:4D	Main effect 2D:4D	0.831	0.365	0.123	0.013	0.146
	2D:4D x Task interaction	2.446	0.123	0.105	0.037	0.338
	2D:4D x Sex interaction	0.743	0.392	0.123	0.012	0.136
	2D:4D x Task x Sex interaction	0.461	0.5	0.105	0.007	0.103

6.5 Discussion

The current experiment sought to explore any potential relationships between 2D:4D, and aspects of 'core' and basic numerical skill. Performances on subitizing, counting, number comparison and SNARC tasks were assessed.

The second (index) to fourth (ring) finger length ratio (2D:4D) has long been identified as a sexually dimorphic anatomical trait (see chapter 1), while a large number of studies have now reported significantly lower 2D:4D in males as compared to females, such sex differences were not replicated in the current study. Male-female differences in 2D:4D however were approaching significance in the predicted direction. Consistent with evidence that the sexual dimorphism in 2D:4D may be greater on the right hand than the left (Manning et al., 1998; Williams et al., 2000, see chapter 2) a stronger sex difference in 2D:4D was observed for right hand 2D:4D measures in comparison to left hand measures. Such findings are also in line with the findings of experiment 2 and 3 where a sex difference in 2D:4D was only revealed for the right hand.

Findings from the current study revealed a significant interaction between left hand 2D:4D and task on percentage error scores, implying a possible relationship between 2D:4D and subitizing task performance relative to control task performance. Low 2D:4D (high PT) participants demonstrated higher accuracy on the control task and reduced accuracy on the subitizing task in comparison to high 2D:4D (low PT) participants.

Interestingly significant task x 2D:4D interaction effects were also revealed in the analysis of counting and number comparison data sets. On the counting task a significant interaction was revealed between task and right hand 2D:4D on percentage error scores and between task and left hand 2D:4D on reaction times. Crucially however the two separate significant interactions implied different relationships between two 2D:4D and performance. The significant interaction between task and right hand 2D:4D on percentage error scores showed greater accuracy in low 2D:4D (high PT) participants on the counting task and greater accuracy in high 2D:4D (low PT) participants on the control task. In contrast, the significant interaction effect between task and left hand 2D:4D on reaction times showed faster reaction times in high 2D:4D participants on the counting task and low 2D:4D participants on the control task. With regard to the significant interaction between right hand 2D:4D and counting percentage error scores however, patterns of reaction time relating to this interaction, although not significant, were in the opposite direction. A possible influence of speed accuracy trade-off effects on this interaction therefore cannot be dismissed. Analysis of number comparison data revealed a significant task x right hand 2D:4D interaction on reaction times, with high 2D:4D participants demonstrating faster reaction times than low 2D:4D participants on the number comparison task and slower reaction times on the control task.

It is interesting to note that there is some degree of consistency in the effects described above in that analysis of subitizing, counting, and number comparison all revealed a significant interaction whereby high 2D:4D is associated with improved performance on the numerical task and low 2D:4D is associated with improved performance on the control task. To the extent that 2D:4D reflects level of PT exposure, such results may imply a detrimental effect of prenatal exposure on basic numerical performance relative to control. Difficult to explain is the lack of consistency in observed effects relating to right vs. left hand 2D:4D measures. Despite the interaction between task and left hand 2D:4D on subitizing percentage error scores for example, the interaction between task and right hand 2D:4D on subitizing percentage error scores was not significant or approaching significance. Similarly with regard to the interactions between task and right hand 2D:4D on counting percentage error scores, task and left hand 2D:4D on counting reaction times, and task and right hand 2D:4D on number comparison reaction times, the same interactions were not significant or approaching significance for the factor of digit ratio derived from the opposite hand.

It is also not the case that significant effects are consistently found for one hand only as significant effects have been demonstrated that relate to both right and left hand

2D:4D measures. These findings are similar to the results of experiment 3 where different interactions involving 2D:4D were identified for analysis including the factor of right hand 2D:4D and analysis including the factor of left hand 2D:4D. Similar to the results of the current study, the interaction effects that were found to be significant were not replicated or even approaching significant when considered in relation to the opposite hand. These inconsistencies raise important questions as to the nature of 2D:4D measures and their adoption as a proxy measure of PT exposure.

It is also important to note that for all of the above described interactions, subsequent post-hoc analysis did not reveal a significant main effect of 2D:4D on subitizing, counting or number comparison task performance (although 2D:4D group differences were approaching significance on the subitizing task). The findings therefore purely imply a possible difference in the relationship between PT exposure and performance on subitizing, counting a number comparison tasks versus a comparable control task.

In line with the findings reported above, correlation analysis revealed significant negative relationships between left hand 2D:4D and counting task reaction times in females suggesting improved performance in high 2D:4D (low PT) participants. While non significant however opposite effects were observed for percentage error scores. It is possible therefore that revealed associations may have been distorted by speed accuracy effects. No further significant simple correlations were found between 2D:4D and performance on any of the numerical tasks.

While experiment 2 and 3 presented no evidence for a direct relationship between 2D:4D and subitizing, the findings did present evidence for a relationship between 2D:4D and lateralization for the process of subitizing relative to a comparable speeded response task. In line with these findings the current study also revealed a significant three way interaction between left hand 2D:4D x task and visual field of stimulus presentation. On the subitizing task, low 2D:4D participants showed a left visual field advantage for subitizing while high 2D:4D participants showed a right visual field advantage. Visual field differences were more pronounced in low 2D:4D participants. Similar to the three way interactions involving a factor of 2D:4D in experiments 2 and 3 however, when this analysis was broken down by task (subitizing vs. alternative) the results revealed no significant two way interactions between 2D:4D and visual field. Furthermore, on inspection of the nature of the significant interactions there is no clear pattern of results across all three experiments.

Bull and Benson (2006) report a stronger SNARC effect in individuals with lower 2D:4D (more masculine) ratios as compared to higher (more feminine) 2D:4D ratios. Contrary to the findings of Bull and Benson (2006) analysis in the current study revealed no significant interaction effects between 2D:4D and task thus no evidence for a significant relationship between 2D:4D and the SNARC effect relative to control. The discrepancy in the findings cannot be explained with reference to differences in the method of analysis between the two studies as subsequent re-analysis of the current results according to the same procedure as that adopted by Bull and Benson (2006) (see Appendix 13) also failed to replicate the findings.

There is some evidence suggesting a certain degree of functional lateralization for number comparison. Research suggests for example, that parietal activation may be greater in the right hemisphere during a number comparison task (Chochon et al., 1999; Stanescu-Cosson et al., 2000). The current study also found a task x visual field interaction for reaction time on the number comparison task whereby, reaction times on the control task were faster for information presented to the RVF (left hemisphere) while average reaction times were faster for information presented to the LVF (right hemisphere) on the number comparison task. Such trends concur with evidence for greater right hemispheric activation during number comparison. Again however post hoc analysis of the data showed no significant effect of visual field on number comparison reaction times following Bonferroni correction. It should also be noted that due to the speed of hemispheric crossover, reaction times to different visual field manipulations may not be entirely reflective of brain lateralization.

Power analysis conducted for all considered effects was again found to be low. Similar to previous experiment, while effects sizes for ANOVA analysis was generally very low effect sizes for correlation analysis revealed a number of small to moderate effects. In an attempt to address the issues of power data from experiments 2 and 3 were combined with results from the current study in order to reconsider possible relationships between 2D:4D and subitizing. Similarly data from experiment 2 was combined with counting data so to also re-examine possible relationships between 2D:4D and counting performance. Despite elevated sample sizes however no significant correlations were revealed between 2D:4D and either subitizing or counting performance. Effect sizes for both analyses actually appeared to be reduced relative to those observed for the consideration of simple correlations within each experiment.

Collectively the results of the present and previous studies provide mixed evidence regarding possible relationships between 2D:4D and basic numerical skills.

One potential limitation which may contribute to the inconsistency in the findings is that the experiments conducted in the current programme of research have, up until now, focused exclusively on possible relationships in adults. As described previously, evidence suggests that our ability to subitize and approximately represent numerical quantity may be related to higher and developing mathematical skills. It is possible however that throughout childhood and adulthood ongoing mathematical education and the development of higher mathematical skills may exert a reciprocal effect on basic numerical capabilities thus impacting upon response times and accuracy to such tasks as those included in the current, and previous two experiments. As such factors were not controlled for in any of the three studies conducted thus far, a possible influence of experience and general mathematical aptitude cannot be dismissed.

Regardless of an individual's level of mathematical education and achievement, a further methodological issues arising from a focus on adults in order to explore possible relationships between 2D:4D basic and core numerical ability relates to the fact that the majority of adults will be well practiced in such skills and thus would be expected to be operating at or close to ceiling on tasks designed to evaluate basic and core numerical performance. This issue is further compounded by the fact that all three experiments carried out thus far in the present programme of research were recruited from a very select, presumably well educated, sample of the population, namely university students. The individual variation in basic numerical performance observed in experiments 2 and 3 and the current study therefore may not be particularly meaningful. In children however the effects of increasing education and experience are reduced. Furthermore, particularly at a young age, basic numerical skills are likely to still be developing.

Even core numerical skills, evident in infants and animals, may still show a degree of developmental progression throughout the primary school years. For example, increases in both the subitizing range and speed of response for the process have been observed. In a study exploring reaction times for judgement of numerosity in children aged 7-15 years old Svenson and Sjoberg, (1983) report response times of 100ms and 71ms to arrays of 1-3 dots in 7 and 8 year olds respectively, rising to 1030ms and 600ms respectively in responses to dot arrays ≥ 4. In the same study the authors report reaction times of 51ms and 47ms in 12 and 15 year olds respectively for the enumeration of quantities 1-4 rising to 450ms and 297ms for the enumeration of quantities ≥ 5. Individual differences across a sample of children in performance on tasks assessing basic and core numerical capacities therefore are likely to reflect variations in the development and mastery of such skills. Evaluation of such differences may thus provide a more meaningful platform on which to assess relationships between PT and basic numerical ability. While a number of previous research studies have explored relations between correlates of PT and basic numerical ability in children (e.g. Finegan et al., 1992; Fink et al., 2006), see chapter 1, no research has, as yet, specifically considered relations between PT and core numerical skills in children. It would be of particular interest therefore for future research to examine relationship between 2D:4D and subitizing and number comparison abilities in infants and children.

In summary, to the extent that 2D:4D is a reflection of exposure to PT, trends revealed in the results of the current investigation may potentially imply a detrimental effect of PT on subitizing, counting and number comparison skills. In line with the finding of experiment 2 and 3 the results also imply a possible relationship between 2D:4D and lateralization for the process of subitizing relative to control. The nature of the interaction between 2D:4D, task (subitizing vs. comparable speed response task) and lateralization however shows no consistent pattern across the three experiments. In addition, significant effects relating to 2D:4D revealed in the current and previous two experiments show no consistency across measures of left and right hand 2D:4D. In contrast to the findings of Bull and Benson (2006) results of the current experiment also provide no support for a possible association between 2D:4D and spatial representations of numerical magnitude (SNARC effect). Overall therefore findings of the current study do not present any strong evidence for a clear relationship between 2D:4D and basic numerical skills in adults.

Chapter 7

Experiment 5: 2D:4D and Core Numerical Skills in Children.

7.1 Introduction

As discussed in chapter 6, studies conducted thus far in the present programme of research have focused exclusively on adults recruited from a university student population. As it is expected however that core numerical abilities will be well-developed and practiced in adult participants, particularly those at a university level education, it is possible that little meaning can be extracted via the assessment of individual variation in such skills. Furthermore, while core numerical skills may impact upon higher numerical and mathematical abilities, the ongoing development of more advanced numerical and mathematical competencies may exert a reciprocal effect on the expression of core numerical abilities. In children, particularly young children, although education and experience may still impact upon core numerical skills, it is anticipated that this influence may be minimised in comparison to adults. What's more there is evidence that the ability to subitize shows a degree of developmental progression with regard to both the subitizing range and reaction times to quantities within that range (Svenson & Sjoberg, 1983). An assessment of core numerical skills in children therefore may offer more meaningful data with regard to individual differences in such capabilities. While relationships between 2D:4D and certain basic numerical skills in children have previously been considered (Finegan et al., 1992; Fink et al., 2006) the potential association between core numerical skills and correlates of prenatal testosterone (PT) have yet to be evaluated in a younger population. The current study will thus attempt to explore possible relationships between 2D:4D as a proxy of PT exposure and tasks designed to assess core numerical skills in children.

In order to investigate relations between 2D:4D and core numerical processing similar subitizing and number comparison tasks (and relevant control tasks) to those adopted with adults in experiment 4 will be utilised in the current study. Such tasks are of a basic nature employing a simple and easy to follow procedure, and were thus deemed suitable for use with both adults and children. Previous research including both adult and child participants has successfully adopted similar numerical tasks in order to assess core numerical skills. Similar to experiment 4 stimuli in the task will be lateralized in order to investigate potential relationship between lateralization and 2D:4D on core numerical task performance. In line with the procedure adopted in previous chapters therefore 2D:4D data will be used to categorise participants into low vs. high 2D:4D groups in order to explore potential interaction between 2D:4D and: task (numerical vs. control), sex and visual filed of stimulus presentation.

Based on the findings of experiment 4 it is hypothesised that a negative relationship may exist between 2D:4D and performance on the core numerical tasks, suggesting a facilitative influence of PT on core numerical abilities. It is also expected that any association between 2D:4D will be distinct from any generic relationship that may exist between 2D:4D and general reaction time task performance thus it is hypothesised that for both subitizing and number comparison a significant interaction will exist between 2D:4D (low vs. high) and task (numerical vs. control). Given the results of experiment 2-4 which imply a potential link between 2D:4D and lateralization for performance on a subitizing vs. control task it is also hypotheses that the factors of 2D:4D and task may also interact with visual field of stimulus presentation (left vs. right). In light of previous evidence and the results of experiment 3 it is further hypothesised that any revealed 2D:4D effects may interact with sex such that the direction and/or the strength of any relationships between 2D:4D and numerical performance may differ between males and females.

7.2 Method

7.2.1 Design

The experiment employed a 2 x 2 x 2 x 2 mixed measures, quasi-experimental design. Separate numerical and control tasks were employed in order to explore subitizing and number comparison skills. The study investigated any main or interaction effects of the factors; 2D:4D (low vs. high), sex (male vs. female), task (numerical vs. control) and visual field of stimulus presentation (left vs. right) on performance. Similar to previous experiments in the current thesis associations between right and left hand 2D:4D and numerical task performance were also explored using correlations.

7.2.2 Participants

Eighty-six (38 male; 48 female) children aged between 6-8 years old were (mean age = 7.3) recruited from mainstream Primary and First schools in and around the North East of England to take part in the experiment. Participants were recruited on a voluntary basis, subject to full informed, written school and parental consent. The parents of participating children provided information regarding their child's date of birth, ethnicity, any potential past or present injury to the second or fourth finger, and any known hormonal abnormalities. No children were excluded for these reasons. Adoption of the exclusion criteria employed in experiment 2 (i.e. excluding all participants not of the majority handedness and ethnicity) resulted in a final sample of 65 participants (30 males, 35 females). Data regarding right hand 2D:4D measurement for one female participant could not be obtained due to technical difficulties in retrieving the saved hand scan. SAT scores for 2 female and 2 male participants could also not be obtained. Similar to experiment 4 technical difficulties resulted in uneven sample sizes across task participation. A total of 54 children, 26 males (mean age = 7.57, SD = 0.69) and 28 females (mean age = 7.23, SD = 0.46) completed the subitizing and subitizing control tasks while a total of 45 participants, 19 males (mean age = 7.39, SD = 0.61) and 26 females (mean age = 7.14, SD = 0.43) completed the number comparison and number comparison control tasks. In line with experiment 4, all analyses were conducted entirely separately for the two subsets of main and control tasks.

7.2.3 Second to Fourth Digit Ratio measurement

The same procedure as that adopted in experiment 4 was used to calculate and evaluate the reliability of second and fourth finger ratio values in the current study. Similar to chapter 4, 2^{nd} and 4^{th} finger measurements, taken from printed images, were carried out by two independent raters to ensure inter-measurement repeatability. Intraclass correlation coefficients (r_I) suggested high inter-rater reliability for second (right hand – 0.97; left hand – 0.968) and fourth (right hand – 0.988; left hand – 0.967) finger and 2D:4D (right hand – 0.883; left hand – 0.854) measurements. TEM and rTEM measurements for the second digit were 0.987 and 0.921 (TEM) and 0.975% and

0.918% (rTEM) for the right and left hands respectively. With TEM and rTEM values for the fourth digit calculated at 0.642 and 0.921 (TEM) and 0.605% and 0.87% (rTEM) respectively for the right and left hands. According to the recommendations of Weinberg et al. (2005), these values are well within an acceptable degree precision for second and fourth finger measurements.

Contrary to expectations no significant sex differences were revealed for either right hand, $t_{(62)} = 0.658$; $p = 0.513$, (males = 0.95, SD = 0.03; females = 0.96, SD = 0.03) or left hand, $t_{(63)} = 0.907$; $p = 0.368$, (males = 0.95, SD = 0.3; females = 0.95, SD = 0.03) 2D:4D measures.

In order to facilitate the analysis of potential complex interaction between task (numerical vs. control), visual field of stimuli presentation (left vs. right), sex (males vs. females) and 2D:4D, within sex median splits according to 2D:4D for both the right and left hands were applied separately for subitizing and number comparison task data sets (see table 41 for median 2D:4D values). T-test analysis confirmed that mean 2D:4D values were significantly different between each group, see Appendix 5.

Table 41

Median male and female 2D:4D values for both the right and left hands for each subitizing and number comparison task data sets.

	Right Hand		Left hand	
	Male	Females	Male	Females
Subitizing	0.947	0.966	0.933	0.954
Number Comparison	0.948	0.956	0.935	0.9461

7.2.4 Adapted Edinburgh Handedness Inventory

Direction of handedness was determined using an adapted version of the Edinburgh Handedness Inventory (Oldfield, 1971) modified for use with children (see Appendix 14). All items on the inventory were read aloud to each participating child and responses were recorded by the experimenter. In order to increase the reliability of recorded answers props were utilised so that children could demonstrate their adopted hand for each described action.

7.2.5 Reaction time tasks

Similar to experiment 4 all tasks were created and administered using the experiment generator package Direct RT (Empirisoft software) and conducted on a Toshiba Tecra M1, Intel Centrino processing laptop with a 14.1" SXGA+ screen. Again there was a pause half way through each task in order to offer participants a short break before continuing (approximately 5-10 minutes). Visual stimuli for each task was developed using Microsoft 'Paint'. For all tasks, written instructions appeared on the screen, all instructions however were also read aloud to each participant. Children were then given a chance to ask questions regarding their understanding of the task before beginning. Instructions for all tasks emphasised both speed and accuracy.

7.2.5.1 Subitizing

The procedure, administration and presentation of the subitizing and subitizing control tasks was identical to that detailed in section 6.3.3.1 of chapter 6 with only one exception, namely that the quantity/colour (2, 3 and 4 dot/ red, blue, yellow square) presentation times in the current study were increased from 100ms to 250ms. Research suggests that time taken to subitize may decrease with age, such that average subitizing times of approximately 100ms are reported in children aged 7, increasing to 71ms in children aged 8, 51ms in children aged 12, 47ms in children aged 15 (Svenson & Sjoberg, 1981). The alteration made to presentation times in the current study therefore was implemented in order to avoid potential floor effects in the adopted sample.

7.2.5.2 Number comparison

The procedure, administration and presentation of the number comparison and number comparison control tasks was identical to that outlined in section 6.3.3.3.

7.2.6 SAT scores

With parental consent, 'Key Stage 1 (KS1)' SAT (Standardised Assessment Tests) scores for each participating child were obtained from schools. KS1 SATs are standardised formal assessments undertaken by children within school under the guidance of their teacher towards to end of year 2 (i.e. when children are aged approximately 7 years old). The tests assess skills in reading and writing, combined to form literacy (includes spelling and handwriting) and mathematics/numeracy (including number, shape, space and measurement) and offer a nationally recognised measure of general educational achievement.

7.2.7 Procedure

The study was approved by the School of Psychology & Sport Sciences Ethics Committee, Northumbria University. On receipt of full informed written consent from the Head Teacher, parental consent forms were distributed. Parents giving their consent were also requested to provide information regarding their participating child's date of birth, ethnicity, any known hormonal abnormalities that the child may possess and any past or previous injury to the second or fourth finger of either hand that the child may have encountered. Children were individually assessed in a quiet room. Prior to testing children were offered a verbal explanation giving basic details of the study, and an assurance of their ability to withdraw at any point or refuse participation on any task. Following informed verbal assent from the child, participants completed the four basic numerical and control computerised tasks. To control for possible order effects tasks were completed in a random sequence. Children were requested to lightly place their head on a chin rest positioned approximately 50cm from the centre of the monitor and focus their gaze towards the centre of the screen. Half way through the computerised tasks participants were given a 15min break during which they completed the adapted version of the Edinburgh Handedness Inventory and hand scans were obtained. Testing took approximately 40 minutes per child following which participating children were fully debriefed. With parental permission, KS1 SAT scores were then obtained from the school.

7.3 Results

In line with the procedure adopted in previous chapters, the means and standard deviations of reaction times (RT) to correct responses and percentage error scores were computed for both the subitizing and number comparison tasks and their associated control tasks. Full tables of means for right and left hand 2D:4D values (for each numerical data set) and performance on each of the numerical tasks can be found in Appendix 5.

Kolmogorov-Smirnov analyses revealed that several variables within the subitizing data set violated normality. In the number comparison data set normality was violated for control task percentage error scores. Where relevant therefore non-parametric test were used to explore behavioural characteristics of the data and simple correlations between 2D:4D and performance.

Similar to experiments 2 -4 however, one of the primary points of interest was the potential complex interactions which might exist between 2D:4D, numerical task performance relative to control task and lateralization for certain numerical. As there is no clear non-parametric equivalent to a 2 x 2 x 2 x 2 ANOVA which would allow for the investigation of such interactions therefore ANOVA analysis was again used to evaluate any main or interaction effects of the factors digit ratio (low vs. high), sex (male vs. female), task (numerical vs. control) and visual field of stimuli presentation (left vs. right) on performance. Similar to analyses conducted in experiment 2 and 4 however the results of this analysis should be considered in the context of the possible loss of test efficiency that might exist due to any violations of normality. Separate analysis was conducted for right and left hand 2D:4D data. For each analysis including a factor of 2D:4D, any significant main or interaction effects relating to 2D:4D are be reported first.

7.3.1 Subitizing

7.3.1.1 Behavioural data

Means and standard deviations of RT to correct responses and percentage error scores on the subitizing and subitizing control task can be seen in table 42. Similar to experiment 4 analysis of any potential 2D:4D effects considered both reaction times and percentage error scores.

Table 42

Means and standard deviations for participant (n = 54) reaction times in ms to correct responses and percentage error for subitizing and control tasks overall and for information presented to presented to the left (LVF) and right (RVF) visual field.

	Subitizing			**Control**		
Task	*LVF*	*RVF*	*Overall*	*LVF*	*RVF*	*Overall*
RT (ms); Mean (SD)	752.65 (189.01)	767.36 (198.25)	747.04 (179.22)	588.44 (134.84)	577.46 (143.35)	575.32 (131.24)
% Error; Mean (SD)	15.29 (11.1)	15.24 (10.29)	15.29 (11.1)	4.38 (7.65)	4.09 (6.5)	4.29 (6.6)

Figure 18 shows average increases in RT as a function of numerosity on the subitizing task and on the control task. Wilcoxn signed ranks analysis (Bonferroni corrected for multiple comparisons = 0.025) revealed significant differences between RT in response to the quantities 2 vs. 3 and 3 vs. 4. Significant differences in percentage error were also revealed in response to the quantities 2 vs. 3. No significant differences in percentage error however were revealed in response to the quantities 3 vs. 4 (see table 43). No significant differences in RT or percentage error were identified in response to the various colours on the subitizing control task.

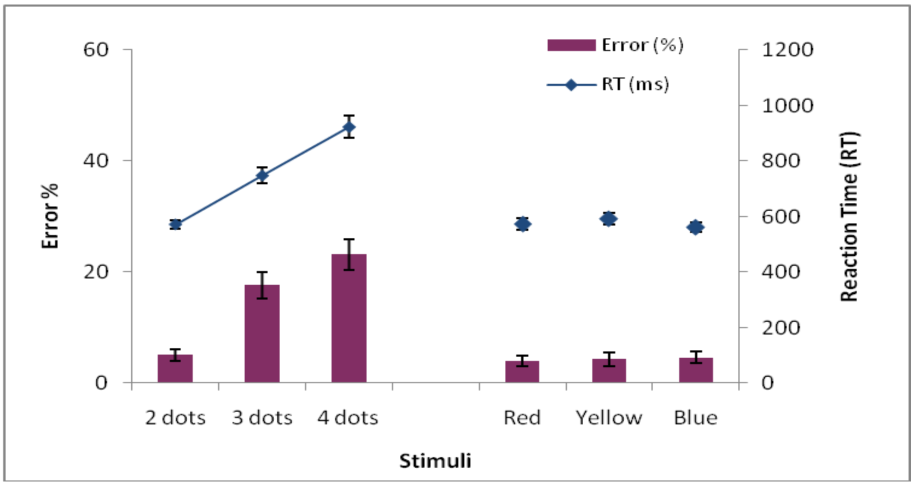


Figure 18. Mean reaction times and percentage error scores for the enumerations of quantities (2-4) in the subitizing task and colours (red, yellow and blue) in the control task, error bars indicate SEM.

Table 43

Average values and Wilcoxn signed ranks analysis of increases reaction time and percentage error as a function of ascending numerosity (n = 54), significant values indicated in bold.

	RT			Percentage error		
Quantity	Mean difference (SD)	Z	p	Mean difference (SD)	Z	p
2-3	176.44	**-6.264**	**<0.001**	12.65	**-4.811**	**<0.001**
3-4	175.34	**-5.889**	**<0.001**	5.5	**-2.134**	**0.033**

No significant speed accuracy associations were identified for performance on either the control task or the subitizing task. Significant negative correlations were observed between age and RTs in responses to both the control, ρ = -0.337; p = 0.013, and subitizing task, ρ = -0.346; p = 0.01, such that older children demonstrated faster response times to the two tasks. No significant correlations were revealed between age and accuracy for either the subitizing or control task.

7.3.1.2 Correlations

As total subitizing reaction times for information presented to both the left and right visual fields and overall (both visual fields combined), and total subitizing percentage error scores for information presented to the right visual field did not meet normality (see Appendix 5) Spearman's correlations were conducted in order to explore any associations between 2D:4D and subitizing performance. As can be seen in tables 44 and 45, while no significant correlations were evident for analysis of all data (males and females combined) significant correlations were identified between left hand 2D:4D and subitizing reaction times for information presented to the left visual field in males, right hand 2D:4D and subitizing reaction times for information presented to the right visual field in females and left hand 2D:4D and subitizing percentage error scores for information presented to the right visual field in females. The correlation between left hand 2D:4D and left visual field subitizing reaction times in males was in a negative direction thus higher 2D:4D values were associated with faster reaction times on the task. While not significant all other correlation relating to possible associations between subitizing reaction times in males and 2D:4D were also in a negative direction. In females however the significant association between right hand 2D:4D and right visual field reaction times was positive, thus lower 2D:4D was associated with faster reaction times. Again, while non significant, the associations between right hand 2D:4D and overall subitizing reaction times and reaction times for subitizing stimuli if females were also in a positive direction. It should be highlighted however that in direct opposition to this finding, the same correlation for percentage error scores, although non-significant, was in a negative direction. Similarly, while the significant correlation between left hand 2D:4D and subitizing percentage error scores for information presented to the right visual field is in a negative direction (thus an association between low 2D:4D and increased error), a contradictory relationship was evident for reaction times, with the same correlation for reaction times revealed to be in a positive direction.

Table 44

Spearman correlation coefficients (ρ) for the relationship between left and right hand 2D:4D and subitizing reaction times. P values and power calculations (1-β) for each analysis are also listed.

	Males and Females n = 54		**Males n = 26**		**Females n = 28**	
Reaction times	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Subitizing Overall	ρ = 0.041 p = 0.767 1-β = 0.059	ρ = -0.16 p = 0.248 1-β = 0.2	ρ = -0.309 p = 0.125 1-β = 0.318	ρ = -0.364 p = 0.068 1-β = 0.428	ρ = 0.315 p = 0.103 1-β = 0.355	ρ = 0.059 p = 0.766 1-β = 0.059
Subitizing Left Visual Field	ρ = 0.025 p = 0.858 1-β = 0.053	ρ = -0.21 p = 0.128 1-β = 0.313	ρ = -0.312 p = 0.212 1-β = 0.323	**ρ = -0.453 p = 0.02 1-β = 0.628**	ρ = 0.288 p = 0.137 1-β = 0.302	ρ = -0.038 p = 0.848 1-β = 0.054
Subitizing Right Visual Field	ρ = 0.151 p = 0.276 1-β = 0.182	ρ = -0.092 p = 0.508 1-β = 0.097	ρ = -0.186 p = 0.362 1-β = 0.14	ρ = -0.246 p = 0.225 1-β = 0.214	**ρ = 0.404 p = 0.033 1-β = 0.553**	ρ = 0.026 p = 0.896 1-β = 0.052

Table 45

Spearman correlation coefficients (ρ) for the relationship between left and right hand 2D:4D and subitizing percentage error scores. P values and power calculations (1-β) for each analysis are also listed.

	Males and Females n = 54		**Males n = 26**		**Females n = 28**	
Percentage error	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Subitizing Overall	ρ = -0.117 p = 0.4 1-β = 0.128	ρ = -0.162 p = 0.241 1-β = 0.203	ρ = 0.11 p = 0.593 1-β = 0.08	ρ = -0.163 p = 0.425 1-β = 0.118	ρ = -0.305 p = 0.114 1-β = 0.335	ρ = -0.184 p = 0.347 1-β = 0.147
Subitizing Left Visual Field	ρ = -0.153 p = 0.27 1-β = 0.186	ρ = -0.188 p = 0.173 1-β = 0.259	ρ = 0.103 p = 0.615 1-β = 0.076	ρ = -0.303 p = 0.132 1-β = 0.307	ρ = -0.29 p = 0.135 1-β = 0.306	ρ = -0.044 p = 0.826 1-β = 0.055
Subitizing Right Visual Field	ρ = -0.087 p = 0.529 1-β = 0.092	ρ = -0.121 p = 0.382 1-β = 0.133	ρ = 0.059 p = 0.776 1-β = 0.058	ρ = 0.019 p = 0.926 1-β = 0.051	ρ = -0.283 p = 0.144 1-β = 0.293	**ρ = -0.378** **p = 0.047** **1-β = 0.492**

7.3.1.3 Right hand 2D:4D, sex and lateralization during subitizing

Analysis revealed a significant two way interaction effect between visual field of stimulus presentation and right hand 2D:4D on reaction times, see table 46. While low 2D:4D (high PT) participants showed faster overall response times to information presented to the RVF (LVF mean = 677.4, SD = 131.09; RVF mean = 655.25, SD = 113.21) the opposite pattern of results was identified for high 2D:4D (low PT) participants who showed faster response times to information presented to the LVF (LVF mean = 663.69, SD = 146.12; RVF mean = 689.57, SD = 173.73). As the factor of 2D:4D did not also interact with task the findings do not reflect a potential association between 2D:4D and lateralization for subitizing but rather a generic effect of 2D:4D on lateralization for both subitizing and control reaction time task. As this is not of specific interest to the current study post hoc analysis of this interaction will not be reported here, the results post hoc analysis of this interaction can be viewed in Appendix 5. As can be seen in tables 46 and 47 no further main or interaction effects involving the factor of right hand 2D:4D on either reaction times or percentage error scores were found to be significant.

Table 46

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of right hand 2D:4D (all df = 1,50) for analysis of subitizing reaction times (significant effect highlighted in bold).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.064	0.801	76150.6	0.001	0.057
2D:4D x Task interaction	0.147	0.703	25568.3	0.003	0.066
2D:4D x Sex interaction	0.319	0.575	76150.6	0.006	0.086
2D:4D x Visual field interaction	**6.787**	**0.012**	4490.12	0.12	0.724
2D:4D x Task x Sex interaction	0.212	0.648	25568.3	0.004	0.074
2D:4D x Visual field x Sex interaction	0.531	0.47	4490.12	0.011	0.11
2D:4D x Task x Visual field interaction	2.645	0.11	7491.6	0.05	0.358
2D:4D x Task x Sex x Visual field interaction	0.186	0.668	7491.6	0.004	0.071

Table 47

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of right hand 2D:4D (all df = 1,50) for analysis of subitizing percentage error scores

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.291	0.592	194.759	0.006	0.083
2D:4D x Task interaction	0.343	0.561	99.233	0.007	0.089
2D:4D x Sex interaction	2.959	0.092	194.759	0.056	0.393
2D:4D x Visual field interaction	0.03	0.864	15.828	0.001	0.053
2D:4D x Task x Sex interaction	3.673	0.061	99.233	0.068	0.468
2D:4D x Visual field x Sex interaction	0.401	0.53	15.828	0.008	0.095
2D:4D x Task x Visual field interaction	0.595	0.444	11.738	0.012	0.118
2D:4D x Task x Sex x Visual field interaction	2.154	0.148	11.738	0.041	0.302

Results of the 4-Way ANOVAs including digit ratio groups split according to right hand 2D:4D measures however did revealed a significant overall main effect of task on both reaction times, $F_{(1,50)} = 65.125$, MSe = 25568.3, $p < 0.001$, $\eta p^2 = 0.566$, 1-β = 1, and percentage error scores, $F_{(1,50)} = 65.573$, MSe = 99.233, $p < 0.001$, $\eta p^2 = 0.567$, 1-β = 1. Significantly faster reaction times and lower percentage error scores were observed for the control task in comparison to the subitizing task (see table 42).

7.3.1.3 Left hand 2D:4D, sex and lateralization during subitizing

4-way ANOVA analysis including the factor of left hand 2D:D revealed a significant three way interaction between left hand 2D:4D, task and visual field on reaction times, see table 48. No further significant main or interaction effect relating to the factor of left hand 2D:4D were revealed.

Table 48

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of left hand 2D:4D (all df = 1,50) for analysis of subitizing and subitizing control task reaction times (significant effect highlighted in bold).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.571	0.453	74985.25	0.011	0.115
2D:4D x Task interaction	1.04	0.313	24717.09	0.02	0.17
2D:4D x Sex interaction	0.65	0.424	74985.25	0.013	0.124
2D:4D x Visual field interaction	0.253	0.617	4958.545	0.005	0.078
2D:4D x Task x Sex interaction	0.963	0.331	24717.09	0.019	0.161
2D:4D x Visual field x Sex interaction	1.834	0.182	4958.545	0.035	0.264
2D:4D x Task x Visual field interaction	**4.115**	**0.048**	6941.642	0.076	0.512
2D:4D x Task x Sex x Visual field interaction	3.106	0.084	6941.642	0.058	0.409

Table 49

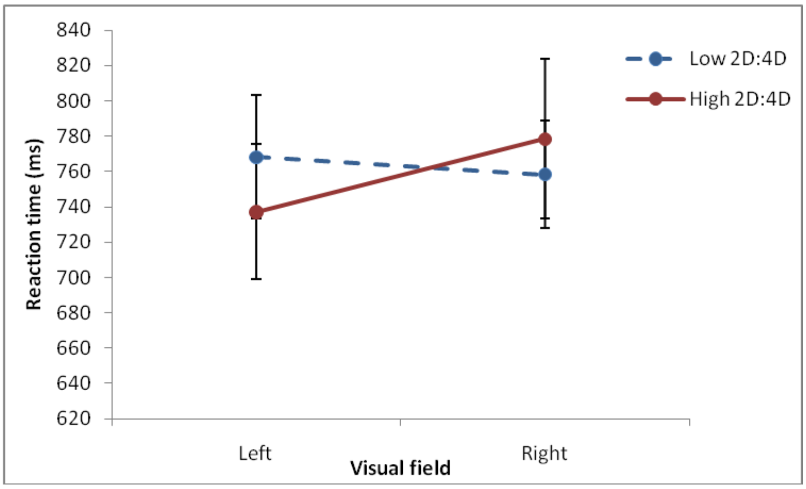
F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of left hand 2D:4D (all df = 1,50) for analysis of subitizing and subitizing control task percentage error scores.

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.332	0.567	203.049	0.007	0.087
2D:4D x Task interaction	0.51	0.478	103.665	0.01	0.108
2D:4D x Sex interaction	0.775	0.383	203.049	0.015	0.139
2D:4D x Visual field interaction	0.002	0.965	15.463	0.00004	0.05
2D:4D x Task x Sex interaction	1.22	0.275	103.665	0.024	0.192
2D:4D x Visual field x Sex interaction	1.622	0.209	15.463	0.031	0.239
2D:4D x Task x Visual field interaction	0.58	0.45	11.926	0.011	0.116
2D:4D x Task x Sex x Visual field interaction	1.492	0.228	11.926	0.029	0.224

As can be seen in figure 19 a and b, on the subitizing task low 2D:4D participants showed a right visual field advantage and faster reaction times than high 2D:4D participants for stimuli presented to the right visual field; while high 2D:4D participants showed a left visual field advantage and faster reaction times than low

2D:4D participants for stimuli presented to the left visual field. This pattern was not replicated for performance on the control task where a right visual field advantage was observed in low 2D:4D participants and a left visual field advantage was displayed in low 2D:4D participants. High 2D:4D participants showed faster reaction times than low 2D:4D participants for information presented to both visual fields. In order to further explore this interaction analysis was split by task and two two-way ANOVAs were conducted to investigate any main or interaction effects of the factors left hand 2D:4D (low vs. high) and visual field (left vs. right) on subitizing and control task reaction times analysed separately. Results of this post-hoc analysis revealed no significant main effect of left hand 2D:4D (subitizing – $F_{(1,52)} = 0.008$, MSe = 67215.654, $p = 0.929$, $\eta p^2 = 0.0002$, $1\text{-}\beta = 0.051$, control – $F_{(1,52)} = 1.911$, MSe = 34860.506, $p = 0.173$, $\eta p^2 = 0.035$, $1\text{-}\beta = 0.274$), no significant left hand 2D:4D x visual field interaction effect (subitizing – $F_{(1,52)} = 2.155$, MSe = 8877.282 $p = 0.148$, $\eta p^2 = 0.04$, $1\text{-}\beta = 0.302$, control – $F_{(1,52)} = 2.742$, MSe = 316.852, $p = 0.104$, $\eta p^2 = 0.05$, $1\text{-}\beta = 0.369$) and no significant main effect of visual field (subitizing – $F_{(1,52)} = 0.658$, MSe = 8877.282, $p = 0.421$, $\eta p^2 = 0.012$, $1\text{-}\beta = 0.125$, control – $F_{(1,52)} = 1.028$, MSe = 3166.852, $p = 0.315$, $\eta p^2 = 0.019$, $1\text{-}\beta = 0.169$) in analysis of either subitizing or control task reaction times.

a)



b)

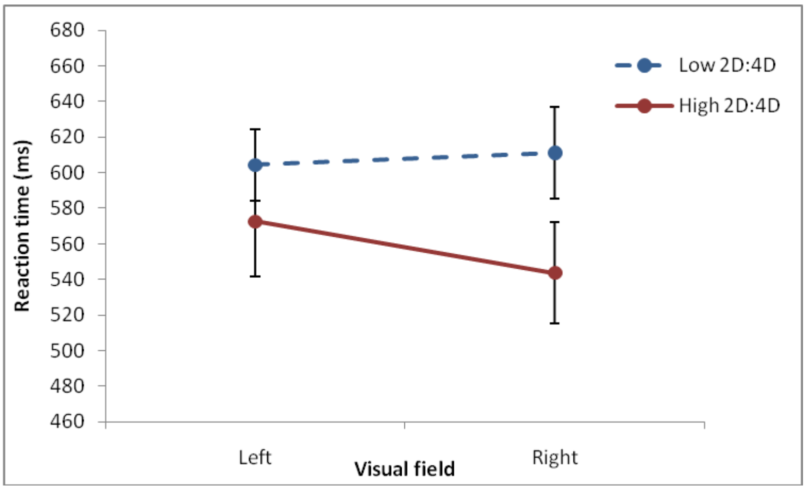


Figure 19 a and b. Mean subitizing (a) and control task (b) reaction times in both low and high 2D:4D participants to stimuli presented to both the left and right visual fields.

4-Way ANOVA analysis including digit ratio groups split according to left hand 2D:4D measures also revealed a significant overall main effect of task on both reaction times, $F_{(1,50)} = 67.368$, MSe = 24717.09, $p < 0.001$, $\eta p^2 = 0.574$, $1\text{-}\beta = 1$, and percentage error scores, $F_{(1,50)} = 62.769$; $p < 0.001$, $\eta p^2 = 103.665$, $1\text{-}\beta = 1$, similar to that described for analysis including right hand 2D:4D measures.

7.3.2 Number Comparison

7.3.2.1 Behavioural data

As percentage error scores on the control task failed to meet the assumption of normality (see Appendix 5) all analyses involving control task percentage error scores were conducted using non-parametric tests. As all other variable of consideration in the number comparison dataset however did not violate the assumption of normality parametric test where adopted reaction time data and number comparison percentage error scores. Table 50 shows means and standard deviations for reaction times to correct responses and percentage error scores on the number comparison and number comparison control task. Analysis of any potential 2D:4D effects considered both reaction times and percentage error scores.

Significant speed accuracy associations were revealed for number comparison task performance overall, $r = -0.586$; $p < 0.001$, and for responses to stimuli presented to the right visual field, r = -0.617; $p < 0.001$, and left visual field, r = -0.422; p = 0.004, analysed separately. All correlations were in a negative direction (percentage error decreased with increasing reaction time) suggesting the presence speed-accuracy trade-off effects. No significant speed accuracy associations however were revealed for performance on the control task. No significant correlations were found between age and response times or percentage errors scores on the number comparison or number comparison control task.

Table 50

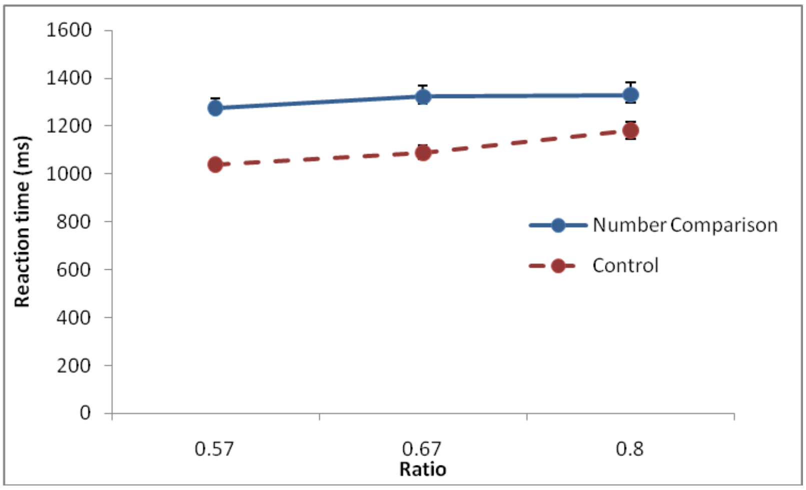
Means and standard deviations for participant (n = 57) reaction times in ms to correct responses and percentage error for the number comparison and control tasks overall and for information presented to presented to the left (LVF) and right (RVF) visual fields

	Number Comparison			**Control**		
	Lvf	*Rvf*	*Overall*	*Lvf*	*Rvf*	*Overall*
RT (ms) Mean	1304.1	1267.7	1287.49	1102.07	1071.01	1087.02
(SD)	(271.11)	(297.48)	(273.98)	(185.15)	(179.87)	(180.06)
% Error Mean	32.83	34.61	33.73	6.56	8.58	7.46
(SD)	(14.35)	(14.53)	(13.2)	(7.11)	(13.45)	(9.17)

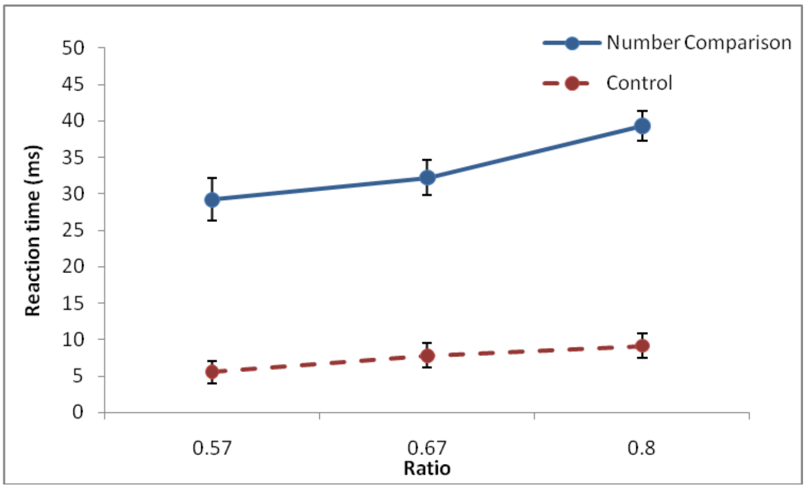
As can be seen in figure 20 a and b, in line with anticipated distance effects, accuracy and reaction time on the number comparison task systematically decreased as the ratio between the two numerosities being compared increased. T-tests analysis however (Bonferroni corrected for multiple comparisons) revealed no significant differences in reaction times to 0.57 vs. 0.67 ratio comparisons, $t_{(44)} = -1.523$; p = 0.135, reaction times to 0.67 vs. 0.8 ratio comparison, $t_{(44)} = -0.246$; p = 0.807, or in percentage error scores for 0.57 vs. 0.67 comparisons, $t_{(44)} = -1.229$; p = 0.225. Significant differences were revealed in percentage error scores between 0.67 vs. 0.8 ratio comparisons, $t_{(44)} = 3.209$; p = 0.002.

With regard to the control task, significant differences were observed between; reaction times to responses for 0.57 ratio comparisons vs. 0.67 ratio comparisons, $t_{(44)} = 2.999$; p = 0.004, reaction times to responses for 0.67 ratio comparisons vs. 0.8 ratio comparisons, $t_{(44)} = 4.664$; $p < 0.001$, and percentage error scores for 0.57 ratio comparisons vs. 0.67 ratio comparisons, Z = -2.257, p = 0.024. No significant differences were revealed in for 0.67 ratio comparisons vs. 0.8 ratio comparisons on the control task, Z = -0.588, p = 0.556.

a)



b)



20 a and b. Mean reaction times for correct responses (ms) and percentage error scores for performance on the number comparison and control task over each ratio level of difference between number and size comparisons, including error bars indicating SEM.

7.2.3.2 Correlations

Tables 51 and 52 show results of the Pearson's correlation analyses conducted to investigate any possible associations between right and left hand 2D:4D and performance on the number comparison task. The analysis revealed no significant associations between measures of 2D:4D and number comparison reaction times and percentage error scores for analysis of all data or male and female data considered separately.

Table 51

Pearson correlation coefficients (r) *for the relationship between left and right hand 2D:4D and number comparison reaction times. P values and power calculations (1-*β*) for each analysis are also listed.*

	Males and Females n = 45		**Males n = 19**		**Females n = 25**	
Reaction times	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
No. Comp. Overall	r = -0.021 p = 0.89 1-β = 0.052	r = -0.014 p = 0.926 1-β = 0.051	r = -0.032 p = 0.896 1-β = 0.052	r = 0.125 p = 0.609 1-β = 0.08	r = -0.025 p = 0.907 1-β = 0.052	r = -0.215 p = 0.291 1-β = 0.179
No. Comp. Left Visual Field	r = -0.038 p = 0.808 1-β = 0.057	r = -0.033 p = 0.829 1-β = 0.055	r = 0.028 p = 0.91 1-β = 0.051	r = 0.141 p = 0.565 1-β = 0.088	r = -0.125 p = 0.553 1-β = 0.091	r = -0.274 p = 0.175 1-β = 0.267
No. Comp. Right Visual Field	r = -0.031 p = 0.844 1-β = 0.055	r = -0.007 p = 0.961 1-β = 0.05	r = -0.042 p = 0.866 1-β = 0.053	r = 0.103 p = 0.674 1-β = 0.07	r = -0.024 p = 0.91 1-β = 0.051	r = -0.163 p = 0.427 1-β = 0.121

Table 52

Pearson correlation coefficients (r) *for the relationship between left and right hand 2D:4D and number comparison percentage error scores. P values and power calculations (1-*β*) for each analysis are also listed.*

	Males and Females n = 45		**Males n = 19**		**Females n = 25**	
Reaction times	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
No. Comp. Overall	r = -0.053 p = 0.732 1-β = 0.064	r = -0.162 p = 0.288 1-β = 0.187	r = -0.072 p = 0.768 1-β = 0.06	r = -0.231 p = 0.341 1-β = 0.158	r = -0.049 p = 0.815 1-β = 0.056	r = -0.104 p = 0.613 1-β = 0.078
No. Comp. Left Visual Field	r = -0.1 p = 0.519 1-β = 0.1	r = -0.237 p = 0.116 1-β = 0.352	r = -0.099 p = 0.686 1-β = 0.068	r = -0.277 p = 0.25 1-β = 0.211	r = -0.113 p = 0.591 1-β = 0.083	r = -0.211 p = 0.302 1-β = 0.174
No. Comp. Right Visual Field	r = 0.006 p = 0.968 1-β = 0.05	r = -0.048 p = 0.754 1-β = 0.061	r = -0.018 p = 0.941 1-β = 0.051	r = -0.126 p = 0.608 1-β = 0.08	r = 0.02 p = 0.925 1-β = 0.051	r = 0.027 p = 0.897 1-β = 0.052

7.2.3.3 Right hand 2D:4D, sex and lateralization during number comparison

No significant main effect of right hand 2D:4D or significant interaction effect involving the factor of right hand 2D:4D were revealed for analysis of either reaction times or percentage error scores on the number comparison and number comparison control tasks, see table 53 and 54.

Table 53

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of right hand 2D:4D (all df = 1,40) for analysis of number comparison and number comparison control task reaction times.

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.002	0.964	171670.7	0.00005	0.05
2D:4D x Task interaction	1.418	0.241	51570.73	0.034	0.213
2D:4D x Sex interaction	0.427	0.517	171670.7	0.011	0.098
2D:4D x Visual field interaction	2.806	0.102	8640.292	0.066	0.373
2D:4D x Task x Sex interaction	0.489	0.488	51570.73	0.012	0.105
2D:4D x Visual field x Sex interaction	0.167	0.685	8640.292	0.004	0.068
2D:4D x Task x Visual field interaction	0.524	0.473	7086.767	0.013	0.109
2D:4D x Task x Sex x Visual field interaction	0.022	0.883	7086.767	0.001	0.052

Table 54

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of right hand 2D:4D (all df = 1,40) for analysis of number comparison and number comparison control task percentage error scores.

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.424	0.519	353.889	0.01	0.097
2D:4D x Task interaction	0.891	0.351	183.715	0.022	0.151
2D:4D x Sex interaction	0.168	0.684	353.889	0.004	0.069
2D:4D x Visual field interaction	0.489	0.489	66.873	0.012	0.105
2D:4D x Task x Sex interaction	1.148	0.29	183.715	0.028	0.182
2D:4D x Visual field x Sex interaction	2.566	0.117	66.873	0.06	0.346
2D:4D x Task x Visual field interaction	2.229	0.143	70.294	0.053	0.308
2D:4D x Task x Sex x Visual field interaction	0.332	0.567	70.294	0.008	0.087

4-way ANOVA analysis including the factor of right hand 2D:4D did reveal a significant main effect of task on both reaction times, $F_{(1,40)} = 36.625$, MSe = 51570.73, $p < 0.001$, $\eta p^2 = 0.132$, $1\text{-}\beta = 0.674$, and percentage error scores, $F_{(1,40)} = 152.961$, MSe = 183.715, $p < 0.001$, $\eta p^2 = 0.793$, $1\text{-}\beta = 1$, with faster response times and higher accuracy observed on the number comparison task, see table 50. A significant overall main effect of visual field was also revealed for analysis of reaction times, $F_{(1,40)} = 6.109$, MSe = 8640.292, $p = 0.018$, $\eta p^2 = 0.132$, $1\text{-}\beta = 0.674$, with faster reaction times displayed for responses to the left visual field (mean = 1203.08, SD = 201.77) as compared to the right visual field (mean = 1169.36, SD = 211.96). Although not significant, percentage errors scores were also lower for information presented to the left visual field, (mean = 19.7, SD = 8.8) as compared to right, (mean = 21.6, SD = 11), the significant main effect of visual field therefore cannot be explained with reference to speed-accuracy trade-off effects.

7.2.3.2 Left hand 2D:4D, sex and lateralization during number comparison

4 way ANOVA analysis including the factor of left hand 2D:4D revealed a significant three way interaction between the factors, left hand 2D:4D, sex and visual field on percentage error scores and a significant three way interaction between the factors left hand 2D:4D, task and visual field also on percentage error scores, see table 56. As can be seen in tables 55 and 56, no further significant main or interaction effects involving the factor of left hand 2D:4D were revealed for analysis of either reaction times or percentage error scores.

Table 54

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of left hand 2D:4D (all df = 1,41) for analysis of number comparison and number comparison control task reaction times.

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.001	0.972	164235.7	0.00003	0.05
2D:4D x Task interaction	0.194	0.662	51391.26	0.005	0.071
2D:4D x Sex interaction	1.299	0.261	164235.7	0.031	0.2
2D:4D x Visual field interaction	0.32	0.574	9274.136	0.008	0.086
2D:4D x Task x Sex interaction	0.751	0.391	51391.26	0.018	0.135
2D:4D x Visual field x Sex interaction	0.284	0.597	9274.136	0.007	0.082
2D:4D x Task x Visual field interaction	0.139	0.771	6808.208	0.003	0.065
2D:4D x Task x Sex x Visual field interaction	1.206	0.278	6808.208	0.029	0.189

Table 55

F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values relating to main and interaction effects for the factor of left hand 2D:4D (all df = 1,41) for analysis of number comparison and number comparison control task percentage error scores (significant effects highlighted in bold).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	0.84	0.365	342.74	0.02	0.146
2D:4D x Task interaction	0.194	0.662	186.679	0.005	0.071
2D:4D x Sex interaction	0.334	0.566	342.74	0.008	0.087
2D:4D x Visual field interaction	0.858	0.36	61.997	0.021	0.148
2D:4D x Task x Sex interaction	0.017	0.895	186.679	0.0004	0.052
2D:4D x Visual field x Sex interaction	**4.853**	**0.033**	61.997	0.106	0.576
2D:4D x Task x Visual field interaction	**4.412**	**0.042**	66.97	0.097	0.536
2D:4D x Task x Sex x Visual field interaction	0.168	0.684	66.697	0.004	0.069

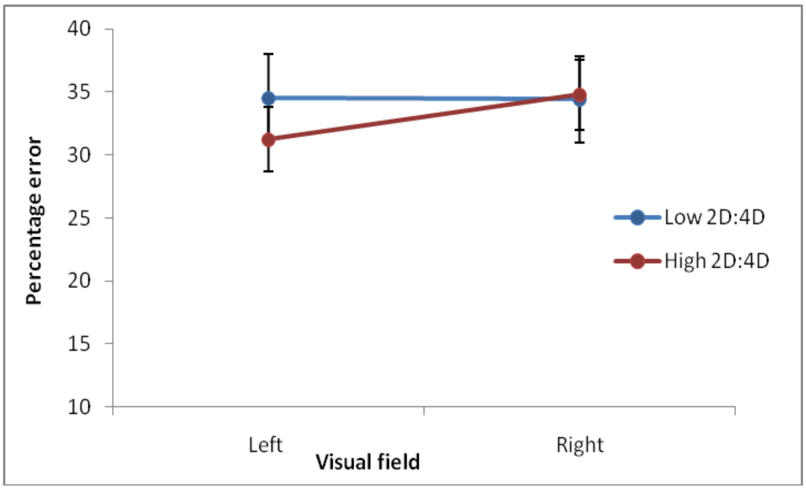
With regard to the three way interaction between left hand 2D:4D, visual field and sex, in males high 2D:4D (low PT) males showed lower percentage error scores than low 2D:4D males for information presented to both the left (mean high 2D:4D males = 20.31, SD = 8.3, mean low 2D:4D males = 20.76, SD = 12.6) and right visual fields (mean high 2D:4D males = 19.36, SD = 8.87, mean low 2D:4D males = 27.26, SD = 19.05) this advantage however was more pronounced for information presented to the right visual field. In addition high 2D:4D males showed greater accuracy for information presented to the right visual field as compared to the left while low 2D:4D males showed greater accuracy for information presented to the left visual field. A dissimilar pattern was displayed in females, where high 2D:4D females showed lower percentage error scores than low 2D:4D participants for information presented to the left visual field (mean high 2D:4D males = 17.86, SD = 7.05, mean low 2D:4D males = 20.33, SD = 8.44) but higher percentage error scores for information presented to the right visual field (mean high 2D:4D females = 20.79, SD = 7.7, mean low 2D:4D females = 20.22, SD = 7.27). 2D:4D group differences in females were also more pronounced for responses to stimuli presented to the left visual field.

In direct opposition to the pattern displayed in males low 2D:4D females showed a slight right visual field advantage while high 2D:4D females showed a left visual field advantage. As the interaction did not include the factor of task the effect potentially reflects a generic relationship between 2D:4D and lateralization that may differ between males and females. As the effect does not present any implications regarding a possible relationship between 2D:4D and numerical skill, and any general relationship between 2D:4D and lateralization is not of specific interest to the current study, further analysis of this interaction will not be reported here. Post-hoc analysis of this interaction however were conducted in which analysis was separated by sex and two 2-way ANOVAs conducted in order to explore any main or interaction effects of visual field and left hand 2D:4D in male and female data analysed separately. Results of this post-hoc analysis and the corresponding reaction time data relating to this analysis can be viewed in Appendix 5.

With reference to the three way interaction between left hand 2D:4D, task (number comparison vs. control) and visual field on percentage error scores, as can been seen in figure 21 a and b low and high 2D:4D participants showed different patterns of visual field preference on both the number comparison and number comparison control task. On the number comparison task low 2D:4D participants showed slightly lower percentage error scores for information presented to the right visual field, while high 2D:4D participants showed lower percentage error scores for information presented to the left visual field. Patterns of percentage error on the control task however where in the opposite direction with low 2D:4D participants demonstrating lower percentage error scores for information presented to the left visual field, and high 2D:4D participants demonstrating lower percentage error scores for information presented to the right visual field. On the number comparison task high 2D:4D individuals displayed greater accuracy than low 2D:4D individuals for information presented to the left visual field, but similar percentage error scores for information presented to the right visual field. On the control task however high 2D:4D displayed greater accuracy than low 2D:4D for information presented to the right visual field and similar percentage error scores to low 2D:4D participants for information presented to the right visual field. Although not significant, patterns of reaction time for this interaction were not identical to patterns of percentage error data. On the number comparison task both high and low participants showed faster reaction times for information presented to the right visual field (low 2D:4D mean = 1276.14, SD = 355.08, high 2D:4D mean = 1259.63, SD = 237.69) as compared to the left visual field (low 2D:4D mean = 1325.52, SD = 341.44, high 2D:4D mean = 1283.61, SD = 186.61) with high 2D:4D participants demonstrating faster reaction times than low 2D:4D participants for information presented to both visual fields. On the control task both high and low participants showed faster reaction times reaction times for information presented to the left visual field (low 2D:4D mean = 1101.32, SD = 184.97, high 2D:4D mean = 1102.78, SD = 189.48) as compared to the right (low 2D:4D mean = 1063.25, SD = 177.69, high 2D:4D mean = 1078.43, SD = 185.61) and low 2D:4D participants demonstrated faster reaction times than high 2D:4D participants for information presented to both visual fields. As the pattern of reaction time data was slightly different to the pattern of percentage error data the influence of possible speed-accuracy trade-off effect on the significant interaction between left hand 2D:4D group, task and visual field cannot be entirely dismissed.

In order to break down the 3-way interaction between left hand 2D:4D group, task and visual field on percentage error scores analysis was split by task and two 2-way ANOVAs were conducted to explore any main or interaction effects of left hand 2D:4D and visual field on percentage error scores on the number comparison and number comparison control tasks analysed separately. No significant main effects of either left hand 2D:4D or visual field were found for analysis of either the control task (left hand 2D:4D – $F_{(1,43)} = 1.274$, MSe = 163.961, $p = 0.265$, $\eta p^2 = 0.029$, $1\text{-}\beta = 0.197$, visual field – $F_{(1,43)} = 1.562$, MSe = 62.853, $p = 0.218$, $\eta p^2 = 0.035$, $1\text{-}\beta = 0.231$), or the number comparison task (left hand 2D:4D – $F_{(1,43)} = 0.135$, MSe = 356.267, $p = 0.715$, $\eta p^2 = 0.003$, $1\text{-}\beta = 0.065$, visual field – $F_{(1,43)} = 1.007$, 67.644, $p = 0.321$, $\eta p^2 = 0.023$, $1\text{-}\beta = 0.166$). No significant left hand 2D:4D x visual field interaction effects were revealed for analysis of either control task ($F_{(1,43)} = 3.61$, MSe = 62.853, $p = 0.064$, $\eta p^2 = 0.077$, $1\text{-}\beta = 0.459$) or number comparison task ($F_{(1,43)} = 1.109$, MSe = 67.644, $p = 0.298$, $\eta p^2 = 0.025$, $1\text{-}\beta = 0.177$) percentage error scores.

a)



b)

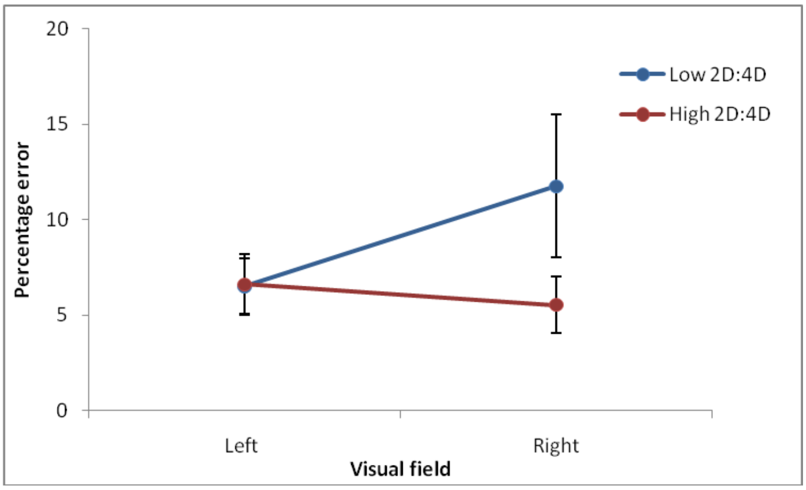


Figure 21 a and b. Mean subitizing (a) and control task (b) reaction times in both low and high 2D:4D participants to stimuli presented to both the left and right visual fields.

Similar to analysis including right hand 2D:4D data 4-way ANOVA analysis including the factor of left hand 2D:4D also revealed a significant main effect of task on both reaction times, $F_{(1,41)} = 35.775$, MSe = 51391.26, $p < 0.001$, $\eta p^2 = 0.466$, $1\text{-}\beta = 1$, and percentage error scores, $F_{(1,41)} = 154.882$, MSe = 186.679, $p < 0.001$, $\eta p^2 = 0.791$, $1\text{-}\beta = 1$, (faster reaction times and greater accuracy on the control task), and a significant overall main effect of visual field (left visual field advantage) on reaction times, $F_{(1,41)} = 4.744$, MSe = 9274.136, $p = 0.035$, $\eta p^2 = 0.104$, $1\text{-}\beta = 0.566$.

7.4 Discussion

The aim of the current experiment was to examine any relationships which may exist between 2D:4D and so called ‘core’ numerical skills in children.

Similar to the findings of experiment 4 well established sex differences in 2D:4D (see chapter 1) were not found in current data. In experiment 4 however sex difference in 2D:4D were approaching significance in the anticipated direction (males < females). In the current study however sex differences were not approaching significance, average right hand 2D:4D ratios were only very slightly lower in males as compared to females while male and females displayed similar average left hand 2D:4D ratios.

The findings of experiment 4 showed significant interactions between task (numerical vs. control) and 2D:4D for both subitizing and number comparison performance. Such findings however were not replicated in the current study. Simple correlation analysis however did reveal a significant negative association between left hand 2D:4D and subitizing reaction times to information presented to the left visual field in males and right 2D:4D and subitizing reaction times to information presented to the right visual field in females. The findings imply a negative effect of PT in males (high 2D:4D associated with lower reaction times) and a positive effect of PT in females (high 2D:4D associated with higher reaction times). Crucially however contradictory results were revealed for females in analysis of percentage error scores with a negative correlation observed between left hand 2D:4D and subitizing percentage error scores for information presented to the right visual field. Furthermore while the findings for reaction times are in line with previous research which suggests sex dependent relationships between potential measures of PT and mathematical performance (Brosnan, 2008; Fink et al., 2006; Finegan et al., 1992; Kempel et al., 2005), the results are contrary to the findings of Fink et al. (2006) and Brosnan (2008) who report lower 2D:4D to be associated with improved performance in males.

The results of experiment 2-4 have all revealed a significant three way interaction between 2D:4D, task (numerical vs. control) and lateralization (as assessed by either visual field or response hand preferences) on subitizing task performance. Consistent with previous findings the current experiment revealed a significant three way interaction between left hand 2D:4D, task (subitizing vs. control) and visual field on subitizing reaction times. Similar to the findings of experiments 2-4 however, closer inspection of this interaction in the current study did not reveal a significant interaction effect of 2D:4D and visual field on either subitizing or control reaction times analysed separately. Evidence from experiments 2-4 however demonstrates no consistent pattern in the nature of the relationship between 2D:4D, visual field and task (subitizing vs. counting/control). With regard to subitizing performance in the current study, low 2D:4D participants showed a right visual field advantage and faster reaction times as compared to high 2D:4D participants for information presented to the right visual field while high 2D:4D participants showed a left visual field advantage and faster reaction times as compared to low 2D:4D participants for information presented to the left visual field. Once again the pattern of results revealed in the current study does not resemble the relationship between 2D:4D and visual field identified in any of the previous experiments.

In contrast to the findings of experiment 4 the current study also revealed a significant interaction between left hand 2D:4D, task (numerical vs. control) and visual field on number comparison percentage error scores. In line with the pattern of results revealed for subitizing reaction times, low 2D:4D participants showed a right visual field advantage, while high 2D:4D participants showed a left visual field advantage for performance on the number comparison task, visual field differences in low 2D:4D participants however were very small. High 2D:4D participants showed lower percentage error scores to information presented to the left visual field during number comparison while number comparison percentage error scores for information presented to the right visual field were similar in low and high 2D:4D groups. Similar to the analysis of subitizing reaction times however no significant 2D:4D x visual field interaction effects were revealed for either number comparison or control task percentage error scores analysed separately. Again therefore, while the results potentially imply the presence of a different relationship between 2D:4D and visual field on number comparison vs. a comparable control task the finding do not suggest a specific effect of 2D:4D on visual field preferences during number comparison. While there are similarities between the left hand 2D:4D x task x visual field interaction effect revelled for subitizing reaction times scores and the same interaction identified for number comparison percentage error scores, given the lack of consistency in the nature of this effect across the present and previous 3 experiments on subitizing performance and the fact that the same interaction for number comparison was not identified in experiment 4, any attempt to interpret or draw conclusions as to implications of this effect would be premature.

In accord with the findings of experiments 2-4, any significant effects that were revealed in ANOVA analysis involving the factor of 2D:4D were identified for analysis of 2D:4D data from one hand only. Significant effects relating to a possible influence of left hand 2D:4D were not replicated or approaching significance for analysis of right hand 2D:4D, similarly significant effects relating to a possible influence of right hand 2D:4D were not replicated or approaching significance for analysis of left hand 2D:4D.

The current experiment revealed a significant interaction between left hand 2D:4D, sex and visual field on number comparison and control task percentage error scores (averaged scores across both task). The interaction implied a potential sex differences in the relationship between 2D:4D and lateralization on general comparison performance. As the interaction however did not also include the factor of task the finding does not suggest a potential effect on sex on the relationship between 2D:4D and lateralization that is specific to number comparison but a more general potential effect of sex on the relationship between 2D:4D and lateralization for similar speeded response comparison tasks.

This is the first experiment to explore relations between a potential correlate of PT and core numerical skills in children, and importantly, a number of vital methodological considerations should still be taken into consideration. Firstly, based on anecdotal observations of participant testing sessions the tasks adopted in the current experiment may not present the most suitable method for the assessment of cognitive performance in children. From a practical perspective, the use of a voice key in order to record reaction time responses was exceptionally difficult to administer in children. Participating children consistently spoke outside of responding or made noise during the task which interfered with voice key recording and meant that, on occasion, a section of task had to be repeated or that data on some trials did not record properly. Such problems in use of the voice key in children heavily contributed to the loss of data through issues with recording.

The low sample sizes included in the current study once again resulted in low power for both ANOVA and correlation analysis. While effect sizes for ANOVA were generally very low a number of small to moderate effect were shown to evident in correlation analysis. Again however current sample sizes did not meet those required in order to achieve power of 0.8 with small to moderate effects (see chapter 4, p. 93).

A further methodological issue was the fact that, despite the inclusion of breaks, the repetitive nature of the tasks clearly impacted on the attention and concentration of a number of children's taking part in the experiment. As the testing session progressed many of the children required a large amount of encouragement in order to focus on the tasks. While the included control tasks did function to offer a measure of any such generalised attentional and concentration factors which may influence performance, given the extent to which this problem was observed it is still possible that such factors may have impacted on the results.

One final methodological limitation relates to the fact that a distance effect was not observed in analysis of reaction times during the number comparison task. Again, as this is a widely replicated effect, failure to demonstrate its significance in the current study raises questions as to the reliability of the adopted number comparison assessment in the evaluation of the 'core' numerical ability to approximately represent and compare large quantities. The numerical tasks utilised in the current study however are similar to those previously adopted in children of the same or a younger age range. In addition, identified reaction times for subitizing with regard to both overall reaction time and increases in reaction time as a function of increasing set size were comparable to those reported in a similar age range by Svenson and Sjőberg (1983). Percentage error scores for number comparison were similar to those reported in a study adopting a similar procedure in order to assess approximate representations of numerical magnitude (Barth et al., 2005). Unfortunately Barth et al. (2005) do not report reaction time data. Reaction times on the number comparison task however are similar to those reported by Landerl et al., (2005) in 8-9 year olds on a digit comparison task. Thus despite the described methodological limitations concerning the assessment of numerical skills the measures do appear to possess a certain degree of reliability in that the performance measures derived from the assessments of subitizing and number comparison do show consistency with previous reports utilising samples of a similar age range.

The numerical tasks used in the current study were specifically designed in order to assess ‘core’ numerical skills and their relationship to 2D:4D as a proxy of PT exposure. The numerical assessments however do not reflect the kinds of tasks encountered in everyday school environments. In adopting such tasks therefore a certain amount of applied value to the research is lost. Given the issues encountered with regard to the assessment of core numerosity in the present study it may be useful for future research to adopt a standardised measure of numerical competency that offers a more direct reflection of core numerical tasks as they may be encountered in a scholastic context.

In conclusion, the results of the current study revealed significant a negative associations between left hand 2D:4D reaction times to subitizing in the left visual field in males. In female a significant positive correlation was identified between right hand 2D:4D and subitizing reaction time to the right visual field. In contrast however a significant negative correlation between left hand 2D:4D and subitizing percentage error scores in the right visual field was also observed in females. In addition the results of the present study again imply a potential association between 2D:4D and lateralization for subitizing performance versus performance on a comparable control task. The results also revealed a novel association between 2D:4D and lateralization for accuracy during number comparison relative to control. Generally however little consistency can be identified across the results of the present experiment and previous experiments conducted thus far in the current programme of research. Given the inconsistency across the findings of the current and previous experiments any concrete conclusions or interpretations of the results are difficult. Much further research perhaps adopting a more standardised and applied assessment of core numerical skills is needed.

Chapter 8

Experiment 6: 2D:4D and Scores on the Dyscalculia Screener in Children.

8.1 Introduction

The experiments described in the previous four chapters have all attempted to explore possible relationships between the second to fourth finger ratio (2D:4D), as a potential proxy of prenatal testosterone (PT) exposure (see chapter 1), and basic so-called 'core' numerical and mathematical skills in adults and children. The findings of the previous four experiments however have demonstrated little observable consistency. The results provide mixed support for a direct relationship between 2D:4D and 'core' or basic numerical task performance and while evidence from the four previous experiments has suggested a possible link between 2D:4D and lateralization for subitizing and number comparison in comparison to control task performance, given the discrepancy in the nature of this effect across studies it is difficult to make any firm interpretations as to the implications of these findings. It is worth noting however that each study has experienced some form of methodological issue. The objective of the present study therefore is to again explore possible relationships between 2D:4D and numerical performance in children, on this occasion however, utilising a standardised battery of basic numerical tasks aimed to assess aspects of 'core' numerical processing and general mathematical skill.

The numerical assessment used in the current study was the Dyscalculia Screener, a standardized, on screen, computer-controlled research and assessment tool targeted at schools and educational professionals. The Dyscalculia Screener was designed by Butterworth (2003) in order to screen individual pupils aged 6 to 14 for dyscalculic tendencies (see chapter 1 for a definition of dyscalculia) and mathematical difficulties.

As described previously, an expanding body of evidence now suggests that our ability to represent and process numerical information is, at least partially, founded upon an innate and potentially evolved core numerical capacity (see chapter 3). While the current thesis has focused specifically on the view that this ability is reflected in two core systems of numerical knowledge, 1) a system for the precise representation of exact numerosities, termed subitizing and, 2) a system of the approximate representation of large numerosities, alternative models regarding the nature of core numerical capacity have been hypothesised. An alternative view proposed by Butterworth (1999) suggests that we are all born with an innate 'number module' characterized as a highly specialised set of neural circuits, located in the left parietal lobe. According to Butterworth's proposed model, this innate 'number module' uses an internal 'numerosity code' that represents numerosities exactly (exact "fiveness", exact "sixness" etc.) and enables us to 'characterize the world in terms of numerosities' (Butterworth, 1999; p 7). Butterworth (1999) identifies certain abilities that are supposedly embedded in our innate number module namely; 1) an understanding that collections of things have a numerosity, that some manipulations of these sets affect the numerosity (e.g. combining collections, taking sub collections away, etc.) and that one collection has the same numerosity as another, or a greater numerosity, or a smaller numerosity, 2) an understanding that numerical collections need not be of visible things; they can equally be audible things, tactile things, abstract things, 3) an ability to recognise small numerosities (i.e. of collections up to about four objects). The critical difference appears to be that the model proposed by Butterworth (1999) does not identify the ability to 'noisily' represent large approximate magnitudes along an analogically compressed 'mental number line' as a fundamental, innate, numerical skill. In contrast to the 'two core systems of numerical knowledge' approach (described in detail in chapter 3) Butterworth's (1999) 'number module' hypothesis assumes that we have an innate capacity to represent the abstract properties of sets, and represent and order the numerosities of sets exactly. While the two approaches differ with regard to the precise characterization of innate numerical competency, both models converge on the notion that, a) innate representations of numerosity and magnitude are multimodal, b) the ability to exactly represent small numerosities (up to 4) constitutes a core numerical skill, c) the ability to compare quantity reflects an innate numerical competency (in the case of the 'number module' this hypothesis refers to comparisons based on exact representations of numerosity; in the 'two core systems of numerical knowledge approach' this hypothesis refers to comparisons based on approximate representations of numerosity along an analogically compressed mental number line).

Unsurprisingly the Dyscalculia Screener is heavily based around Butterworth's theory of the development of mathematical competence. The software is designed to diagnose developmental dyscalculia using tasks designed to assess individual differences in the function of the hypothetical 'number module'. The tasks adopted in the screener are those that research has identified to be the most effective in discriminating individuals with dyscalculia from other individuals (e.g. Landerl et al., 2004). The Dyscalculia Screener consists of four sub-tests, 1) Simple reaction time, in order to assess baseline speed of response, 2) and 3) Dot enumeration and Number Comparison (in the form of numerical stroop), both tests of basic numerosity (forming the so called capacity subscale), and 4) Arithmetic achievement test, addition problems only for children aged 6-9 years, addition and multiplication for children aged 10 years and above (the so called achievement subscale). All tasks incorporated in the screener are item-timed thus the inclusion of a measure of response speed allows for the identification of children who are able to achieve an average number of correct answers, but solve them in an abnormal or an abnormally slow manner.

The Dyscalculia Screener offers a standardised research tool that is accessible to a broad age range (6-14) and simple to administer. As the tool is designed for application in an educational context any possible relationships identified between 2D:4D and performance on the screener would have particularly important implications with regard to or understanding of factors which may impact a child's numerical/mathematical performance and understanding. In adopting the Dyscalculia Screener in the current study the aim is to re-assess possible associations between 2D:4D and both simple numerical capacities and simple arithmetical achievement at a more general and scholastically applied level without moving away from a focus core, innate numerical skills. As the tasks included in the screener have been standardised for use with young children it is hoped that the methodological issues encountered in the tasks utilised thus far in the present programme of research will be avoided. Based on previous evidence using more standardised measures of mathematical skill (e.g. Brosnan, 2008; Fink et al., 2006, see chapter 1) it was hypothesised that a significant correlations would exist between 2D:4D, as a potential proxy of PT, and performance on the Dyscalculia Screener. Also based on previous evidence that the relationship between 2D:4D and performance on certain mathematical measures may be different in males and females or only significant in one sex (e.g. Finegan et al., 1992; Fink et al., 2006; Kempel et al., 2005) it was hypothesised that the direction and strength of possible associations between 2D:4D and performance on the numerical measures may be different in males and females.

8.2 Method

8.2.1 Design

The study employed a correlational design in order to explore the relationship between 2D:4D and performance on the Dyscalculia Screener in males and females.

8.2.2 Participants

A total of 68 children (40 male; 28 female) with a mean age of 8 years (SD = 0.83) were recruited to partake in the experiment. Participating children were recruited on a voluntary basis from mainstream primary and first schools in and around the North East of England following full informed, written school and parental consent. Based on parental report all participating children were of white British ethnic origin, had acquired no present or previous injury to the second or fourth finger of either hand and possessed no known hormonal abnormalities. In line with the exclusion criteria employed in all previous experiments, all participants not of the majority handedness (right handed as assessed by an adapted version of the Edinburgh Handedness Inventory) were removed from the data. Due to technical difficulties when recording scores on numerical measures, data from 1 male participant was also lost. In total, with exclusion and consideration of missing data the overall sample consisted of 34 males (mean age = 8, SD = 0.72) and 24 females (mean age = 7.89, SD = 1). Within this sample however data regarding right hand 2D:4D measurement for one female participant and left hand 2D:4D data for a different female participant could not be obtained due to technical problems in retrieving the saved hand scans. In addition one male participant was classed as ‘working towards’ level 1 (the lowest level) in Key Stage 1 literacy thus a SAT literacy score could not be recorded for this participant.

8.2.3 Second to Fourth Digit Ratio measurement

The same procedure as that adopted in experiment 4 for measuring and assessing reliability of the second and fourth finger ratio was employed in the current study. Intraclass correlation coefficients (r_I) suggested high inter-rater reliability for second (right hand – 0.994; left hand – 0.986) and fourth (right hand – 0.992; left hand – 0.992) finger and 2D:4D (right hand – 0.965; left hand – 0.931) measurements. TEM and rTEM measurements for the second digit were 0.367 and 0.504 (TEM) and 0.497 and 0.69 (rTEM) for the right and left hands respectively. With TEM and rTEM values for the fourth digit calculated at 0.435 and 0.42 (TEM) and 0.567 and 0.545 (rTEM) respectively for the right and left hands. Again, these values are well within an acceptable degree precision for second and fourth finger measurements (Weinberg et al., 2005).

As anticipated, males demonstrated comparatively lower average 2D:4D ratios (right hand mean = 0.956, SD = 0.036, left hand mean = 0.952, SD = 0.034) than females (right hand mean = 0.972, SD = 0.039, left hand mean = 0.963, SD = 0.033). This sex difference however was not revealed to be significant for either right, $t_{(55)} = 1.584$; $p = 0.119$, or left $t_{(55)} = 1.307$; $p = 0.197$, hand 2D:4D measures.

8.2.4 Adapted Edinburgh Handedness Inventory

The present study employed the same adapted version of the Edinburgh Handedness Inventory (Oldfield, 1971) as that use in experiment 5. The procedure adopted in administration of the inventory was identical to that detailed in chapter 7 (see section 7.2.4).

8.2.5 SAT scores

Also similar to the previous experiment Standardised Assessment Test (SAT) scores for literacy and mathematics were obtained. See chapter 7 section 7.2.6 for details.

8.2.6 Dyscalculia Screener

As noted previously, the Dyscalculia Screener consists of four sub-tests namely, 1) Simple reaction time, 2) Dot Enumeration, 3) Number Comparison, and 4) Arithmetic (administered in this respective order). The three numerical tasks (Dot enumeration, Number comparison and Arithmetic) are divided into two subscales; 1) a capacity subscale (involving the Dot enumeration and Number comparison task) and, 2) an achievement subscale (involving the Arithmetic task). Throughout the test children are given both written and audio instructions. Children are asked to perform the tasks as quickly as possible and both RT (to the nearest millisecond) and error is recorded. Each subtest is preceded with practice trials, which may be repeated if required. According to the specified criteria of the test, low performance on the capacity subscale and achievement subscale reflects a pattern of results evident of dyscalculia. Age appropriate performance on the capacity tests despite low arithmetic achievement however is thought to indicate problems in arithmetic due to either social/experiential factors and/or alternative general or specific cognitive deficits rather than dyscalculia defined as a specific deficit in an innate understanding of numerosity. The screener provides the median reaction time, a standard age score and a stanine score (measure of performance on a standardized nine point scale). For numerical sub tests (dot enumeration, number comparison, arithmetic) a percentage correct score, and so called efficiency measure is also provided. Median RTs are calculated for correct responses only. Standardised age scores are calculated as the median RT adjusted to account for median simple reaction time. Efficiency measures represent this standardised age score divided by the proportion of correct answers. The test takes approximately 20-30 minutes to administer depending on the age and response times of the child.

8.2.6.1 Simple RT

The Simple RT task is designed to measure baseline speed of response and is utilised a covariate in order to adjust recorded RTs on the numerical tests as a function of generic reaction time. During this sub-test children are required respond (pressing a key on either left or right of the screen, depending on trial block) to a single black spot presented on a white background on the computer screen. Dots were presented in random positions over a total of 30 trials (15 trials per each response hand block).

8.2.6.2 Dot enumeration

The second and third trials are referred to by Butterworth (2003) as tests of numerical capacity and are designed to assess basic numerosity. During the Dot Enumeration sub-task children are asked to compare the number of dots (up to 9) on half the screen with the Arabic digit (1-9 inclusive) on the other half of the screen and response yes or no as to whether the two numerosities match (no = left hand response, yes = right hand response). Dots and Arabic digit were coloured white and presented on a white background. In total 68 stimuli are presented in a fixed, pseudo-random order. According to Butterworth (2003) in order to complete this task the child must be able to judge the number of dots in an array, either via subitizing or counting, and must be able to demonstrate an understanding of the meaning of the numerals 1-9. The task aims to assess the capacity to represent exact numerosities and knowledge of the numerosity that each numeral represents (Iuculano et al., 2008).

8.2.6.3 Number comparison

For the number comparison task children are asked to select the numerically larger of two Arabic numerals (black numerals on a white background) one presented to one half of the computer screen and one to the other half (left hand response if left side has largest number, right hand if the right side displays the highest number). The task takes the form of a numerical stroop whereby stimuli varies with regard to both the physical and numerical size of the numerals (equal numbers of congruent and incongruent trials are presented). In total 42 trials were presented in a fixed, quasi-random order. A fluent understanding of numerals, knowledge of the connection between numerals and their meanings and the capacity to order numerals by magnitude are identified as necessary for this task (Butterworth, 2003; Iuculano et al., 2008).

8.2.6.4 Arithmetic

Butterworth (2003) identifies the final task, Arithmetic, as a test of numerical achievement designed to distinguish between children who have already learned arithmetical facts and can retrieve them from memory from those who still have to calculate the answer using their fingers for example. During this task arithmetic problems are presented on the screen with an answer, e.g. 3+5 = 8, and children are asked to respond whether the answer is correct (left hand for no, right hand for yes). For younger children this task consists of only addition, for older children (10 years old and above) multiplication tasks are also presented. As all participating children in the current study were aged under 10 years old multiplication problems did not form part of the assessment for any participants in the present experiment . Single-digit addition problems were written in black and presented on a white background over 28 trials presented in a fixed pseudo-random order.

8.2.7 Procedure

The study was approved by the School of Psychology & Sport Sciences Ethics Committee, Northumbria University. On receipt of full informed written consent from the school's head teacher, parental consent forms were distributed among parents of children aged 6-8 years old in participating schools. Parents giving consent were also requested to provide information regarding their child's date of birth, ethnicity, any known hormonal abnormalities that the child may possess and any past or previous injury to the second or fourth finger of either hand that the child may have encountered. Children were individually assessed in a quiet room. Prior to testing the children were given a verbal explanation providing basic details of the study, and an assurance of their ability to withdraw at any point or refuse participation on any task. Following verbal assent from the child, participants completed the Dyscalculia Screener. Following completion of this test the children then completed the adapted version of the Edinburgh Handedness Inventory and hand scans were obtained. Testing took approximately 35 minutes per child, after which the participants were fully debriefed. With parental permission, key stage 1 SAT scores were obtained from the school.

8.3 Results

8.3.1 Behavioural data

As in previous chapters, Kolmogorov-Smirnov tests were used in order to explore normality. The results of these analyses revealed that scores on the majority of performance measures on the dyscalculia screener were not normally distributed, see Appendix 6, two-tailed Spearman correlation coefficients (ρ) were thus used to assess the relationship between 2D:4D and performance on the test battery. Mann-Whitney *U* tests were adopted in order to analyse sex differences in performance. Table 56 shows means and standard deviations of median RTs, standard age scores, stanine scores, percentage correct and efficiency measure scores on all subtests for the entire data set (males and females combined). Full tables of means for performance on the dyscalculia screener in males and females separately can be viewed in Appendix 6. None of scores for the participating children in the current study met the criteria for a classification of dyscalculia according to the specifications outlined by Butterworth (2003).

Table 56

Means and standard deviations (SD) of participant scores (n = 58) on all performance measures on the dyscalculia screener for all sub-tests.

	Simple RT	Dot Enumeration	Number Comparison	Arithmetic
Median RT	429.18 *(142.36)*	2993.33 *(1168.09)*	1338.96 *(467.85)*	5052.76 *(3271.04)*
Standard Age	93.59 *(12.08)*	104.88 *(15.57)*	106.81 *(16.3)*	96.93 *(14.43)*
Stanine	4.07 *(1.66)*	5.24 *(1.84)*	5.5 *(1.96)*	3.76 *(1.73)*
Percentage correct		85.4 *(10.74)*	92.18 *(16.25)*	77.16 *(16.32)*
Efficiency measure		7044.76 *(1237.08)*	3987.98 *(523.85)*	14247.13 *(3522.62)*

As can be seen from table 57 significant associations were found between age and all measures of performance on the number comparison task excluding standard age scores. On the dot enumerations task correlations with age were significant with all measures except percentage error and standard age scores. Significant correlations were also revealed between age and median RTs on the simple RT task and age and stanine scores on the arithmetic task. Where correlations were significant all implied improved performance (higher scores/ reduced RT or error) with increasing age.

Mann-Whitney *U* tests revealed no significant sex differences for any of the performance measures relating each of the subcategories in the dyscalculia screener. Statistics relating to analysis of sex differences in performance on the test battery can be viewed in Appendix 6.

Table 57

Matrix of Spearman's correlation coefficients (ρ) demonstrating relationships between age (in months) and all performance measures for each sub-test on the dyscalculia screener.

	Simple RT	**Dot Enumeration**	**Number Comparison**	**Arithmetic**
Median RT	-0.306*	-0.515***	-0.474***	-0.12
Standard age score	-0.071	0.245	0.239	0.058
Stanine	-0.069	0.322*	0.352**	0.341**
Percentage correct		-0.041	0.33*	0.204
Efficiency measure		0.531***	0.482***	0.127

* significant at 0.05 level; ** significant at 0.01 level; *** significant at 0.001 level.

8.3.2 Correlations between 2D:4D and performance on the dyscalculia screener

As a total of 108 correlation analyses are required in order to analyse all possible relationships between right and left hand 2D:4D and performance measures for each subcategory of the dyscalculia screener, overall and in males and females separately, only the results relating to median reaction time and percentage correct scores will be reported here. The results of analyses relating to standard age scores, stanine scores and efficiency measure data however can be viewed in Appendix 6. Power for Spearman's was calculated in G*Power 3.1.2 (Faul et al., 2009) using adjusted sample sizes (in order to take into account the loss of power relative to its parametric equivalent) as described in chapter 2, section 2.3.1 (p. 51). Tables 58 and 59 show the results of the Spearman's correlation analysis. While several correlation coefficients suggested small to moderate effect sizes for positive association between 2D:4D and number comparison task performance no significant associations were revealed between right or left hand 2D:4D measures and reaction times or accuracy on any of the subtasks on the Dyscalculia Screener in the either entire data set, or male and female data analysed separately.

Table 58

Spearman's correlation coefficients (ρ), p values and 1-β values for analysis of the relationship between right hand 2D:4D median reaction times and percentage correct scores for all subtasks of the Dyscalculia Screener in males (n =34), females (n =24) and the entire data set (n = 58).

		Simple RT	**Dot enumeration**	**Number comparison**	**Addition**
Median RT	Males	ρ = 0.176 p = 0.32 1-β = 0.157	ρ = 0.079 p = 0.657 1-β = 0.07	ρ = -0.085 p = 0.632 1-β = 0.074	ρ = 0.017 p = 0.926 1-β = 0.051
	Females	ρ = -0.251 p = 0.249 1-β = 0.196	ρ = 0.255 p = 0.24 1-β = 0.202	ρ = -0.17 p = 0.438 1-β = 0.114	ρ = -0.074 p = 0.737 1-β = 0.061
	Overall	ρ = -0.002 p = 0.989 1-β = 0.05	ρ = 0.201 p = 0.134 1-β = 0.3	ρ = -0.055 p =0.683 1-β = 0.067	ρ = -0.006 p = 0.968 1-β = 0.05
Percentage correct	Males		ρ = -0.062 p = 0.726 1-β = 0.063	ρ = 0.188 p = 0.286 1-β = 0.174	ρ = -0.008 p = 0.963 1-β = 0.05
	Females		ρ = 0.074 p = 0.737 1-β = 0.061	ρ = 0.123 p = 0.577 1-β = 0.082	ρ = -0.055 p = 0.803 1-β = 0.056
	Overall		ρ = 0.082 p = 0.543 1-β = 0.089	ρ = 0.208 p = 0.121 1-β = 0.319	ρ = -0.028 p = 0.836 1-β = 0.054

Table 59

Spearman's correlation coefficients (ρ), p values and 1-β values for analysis of the relationship between left hand 2D:4D and median reaction times and percentage correct scores on the Dyscalculia Screener in males (n =34), females (n =24) and the entire data set (n = 58).

		Simple RT	**Dot enumeration**	**Number comparison**	**Addition**
Median RT	Males	ρ = 0.121 p = 0.496 1-β = 0.099	ρ = -0.02 p = 0.911 1-β = 0.051	ρ = 0.019 p = 0.915 1-β = 0.051	ρ = -0.139 p = 0.431 1-β = 0.115
	Females	ρ = -0.088 p = 0.688 1-β = 0.066	ρ = 0.068 p = 0.757 1-β = 0.06	ρ = 0.158 p = 0.471 1-β = 0.105	ρ = -0.018 p = 0.936 1-β = 0.051
	Overall	ρ = 0.072 p = 0.597 1-β = 0.08	ρ = -0.038 p = 0.778 1-β = 0.058	ρ = 0.11 p = 0.415 1-β = 0.121	ρ = -0.083 p = 0.539 1-β = 0.09
Percentage correct	Males		ρ = -0.191 p = 0.279 1-β = 0.178	ρ = 0.178 p = 0.314 1-β = 0.16	ρ = -0.058 p = 0.746 1-β = 0.061
	Females		ρ = 0.01 p = 0.962 1-β = 0.05	ρ = 0.238 p = 0.275 1-β = 0.18	ρ = -0.141 p = 0.522 1-β = 0.093
	Overall		ρ = -0.091 p = 0.502 1-β = 0.098	ρ = 0.217 p = 0.105 1-β = 0.343	ρ = -0.086 p = 0.527 1-β = 0.093

8.4 Discussion

There is evidence for a potential relationship between correlates of PT and aspects of mathematical and numerical processing (e.g. Brosnan, 2008; Finegan et al., 1992; Fink et al, 2006; Kempel et al., 2005; Luxen & Buunk, 2005). Thus far however, the research described this current thesis has provided mixed and inconsistent evidence for a possible relationship between a somatic marker of PT (2D:4D) and so called ‘core’ and basic numerical skills. The aim of the current experiment was to re-assess this possible association utilising a standardized assessment of basic mathematical competence in children, namely the Dyscalculia Screener developed by Butterworth (2003). Findings revealed no significant associations between 2D:4D and performance on the Dyscalculia Screener in overall data or males and females analysed separately.

Despite being in the expected direction (males < females), similar to experiments 4 and 5, data gathered in the current study, did not replicate longstanding evidence (George, 1920; Manning et al., 1998; Manning, 2002; Phelps, 1952) for sex differences in 2D:4D. In line with previous evidence suggesting a lack of sex differences in basic numerical skill (Brosnan, 2008; Bull et al., 2009; Fink et al., 2006; Geary, 1996) no significant sex difference were revealed for any of the performance measures in each subcategory of the Dyscalculia Screener.

The results of the current study in relation to 2D:4D are contrary to experiments 1, 4 and 5 which all revealed some evidence for potential effects linking 2D:4D to at least one aspect of numerical performance. Similar to experiments 2 and 3 therefore the findings are contrary to previous evidence for significant associations between 2D:4D and numerical skill (Brosnan, 2008; Finegan et al., 1992; Fink et al., 2006; Kempel et al., 2005; Luxen & Buunk, 2005). It is important to recognise however that power for all analyses was again very low. As some of the revealed correlation coefficients suggest the presence of small to moderate effect sizes it is possible that larger samples may have resulted in significant effects.

Chapter 9

Experiment 7: Re-assessing the relationship between 2D:4D and National Standard Assessment Test (SAT) scores

9.1 Introduction

Previous experiments in the current thesis have attempted to explore a possible relationship between 2D:4D, as a somatic marker of prenatal testosterone (PT) exposure, and basic and potentially innate numerical skills in children and adults. The experiments conducted thus far however have failed to identify any consistent relationship between 2D:4D and numerical performance in either population. Throughout previous experiments a clear lack of consistency in the nature and/or significance of relationships with 2D:4D have also been observed across different tasks. It is extremely difficult therefore to compare the findings of the previous experiments with previous research suggesting a possible relationship between correlates of PT and aspects of numerical performance based on variety of different numerical and mathematical assessments (e.g. Brosnan, 2008; Finegan et al., 1992; Fink et al, 2006; Kempel et al., 2005; Luxen & Buunk, 2005).

In addition to the assessment of basic numerical skills, experiments 5 and 6 also collected data regarding Key Stage 1 (KS1) Standardised Assessment Task (SAT) Scores in literacy and numeracy which were not analysed (see chapter 7, section 7.2.6 for information on the nature of KS1 SAT assessments). Previous research by Brosnan (2008) investigated the association between 2D:4D measures and KS1 SAT scores in 75 children aged 6-7. Brosnan (2008) found a significant negative correlation between 2D:4D and the difference between numeracy and literacy (SAT numeracy scores minus SAT literacy scores). As low 2D:4D is thought to indicate higher PT exposure, the results implied a facilitative effect of PT on numeracy relative to literacy. When the data was split by sex, Brosnan (2008) reported a significant negative correlation between 2D:4D and SAT numeracy scores and the difference between SAT numeracy and literacy scores, but no significant correlation between 2D:4D and SAT literacy scores in males. In females however a significant positive correlation was identified between 2D:4D and SAT literacy scores while no significant associations were found between 2D:4D and; SAT numeracy scores and the difference between SAT numeracy and literacy scores. The results potentially suggest a relationship between lower PT exposure and enhanced verbal abilities in females and higher PT exposure and superior numerical abilities in males. The findings were in accord with evidence from Fink et al. (2006) who reported significant negative relationships between 2D:4D and number knowledge, counting, and visual number representation in males aged 6-11 in the absence of similar significant associations in females of the same age range. Fink et al. (2006) however did not assess any potential relationships between 2D:4D and verbal ability.

The findings of Brosnan (2008) are interesting, given the applied nature of the measure of numerical performance and the similarity of the results to those described by Fink et al., (2006). Interpretation of the evidence present by Brosnan (2008) however requires a certain degree of caution for a number of reasons. Firstly, the relationships between 2D:4D and SAT performance identified by Brosnan (2008) relate to 2D:4D measured as an average digit ratio calculated as the mean of left and right hand 2D:4D measures. This is not a widely used technique by which to measure 2D:4D. When correlations were broken down by right and left hand 2D:4D data, the only correlation that remained significant was that between left hand 2D:4D and SAT numeracy scores in males. The same correlation with respect to the right hand was not significant or approaching significance ($p = 0.38$). Given that separate analysis including left vs. right hand 2D:4D data have failed to show a consistent pattern of results throughout the experiments conducted in this thesis the justification for collapsing left and right hand 2D:4D values into a single measure of 2D:4D may be questioned.

Secondly, Brosnan (2008) reported that out of the sample 95% of the participants were Caucasian and 10 children predominantly used their left hand for class work. As described in previous chapters, while the magnitude of sex differences in 2D:4D appears to be similar across different populations (Manning, Henzi, 2003; Manning, Stewart, et al., 2004; Peters et al., 2007) population and racial differences have been observed in both adults (Manning, Barley, 2000; Manning, Henzi, 2003; Peters et al., 2002) and children (McIntyre et al., 2005; Manning, Stewart, et al., 2004) and may actually account for a greater proportion of the variation in 2D:4D than sex (McIntyre, 2006). Brosnan (2008) however failed to control for the factor of ethnicity in his sample. Similarly, there is evidence that handedness may be related to PT and, more specifically, 2D:4D (Fink et al., 2004; Manning, Trivers, et al., 2000; Manning & Peters, 2009; Nass et al., 1987; Nicholls et al., 2008, see chapter 1), again however Brosnan (2008) failed to control for this. While Brosnan (2008) did report that there were no statistical differences between digits or performance in left vs. right handed individuals the extent to which average differences would be observed is hugely reduced given that the sample of left handed participants consisted of only 10 participants, it is still possible therefore that minor variations in 2D:4D and performance across left and right hand individuals may have influenced the results.

Finally, Brosnan (2008) presented a directional hypothesis thus all reported p-values were one-tailed. As described in chapter 1 however only a handful of studies to date have actually attempted to directly explore the relationship between PT exposure and mathematical ability and there are a number of inconsistencies in previous findings with regards to the nature of possible relationships. For example, while the results of Brosnan (2008) were in line with the findings of Fink et al., (2006), Finegan, et al. (1992) reported a negative relationship between amniotic fluid testosterone and performance on counting and number fact tasks in girls, such that high testosterone levels were associated with poorer performance. The use of one-tailed analysis therefore may not be entirely justified at this stage and may have inflated the possibility of Type I error.

The current experiment attempted to replicate the findings of Brosnan (2008) utilising the SAT data collected in experiments 5 and 6. Based on the findings of Brosnan (2008) it was hypothesised that 2D:4D would be related to SAT performance and that the nature of any relationship between 2D:4D and performance may be different for SAT numerical and SAT literacy scores. Based on previous evidence it was further hypothesised that the nature and strength of any potential associations between 2D:4D and SAT literacy and numeracy scores may be different for males and females.

9.2 Method

9.2.1 Design and procedure

Data regarding age, sex, 2D:4D and KS 1 SAT numeracy and literacy data were collated from chapters 7 and 8. A correlational design was then used in order to explore the relationship between 2D:4D and SAT numeracy and literacy scores in males and females.

9.2.2 Participants

The combination of data from experiments 7 and 8 resulted in an initial sample of 154 participants. According to the exclusion criteria adopted in previous chapters however participants not of the majority handedness (i.e. right handed, assessed according to an adapted version of the Edinburgh Handedness Inventory, see Appendix 14 and chapter 7 for details) or ethnicity (white British assessed according to parental report) were not included in the analysis. Any participants with missing SAT data were also excluded, resulting in a final sample of 119 right-handed, Caucasian British children aged 6-9 years old. As detailed in chapter 7 and 8, all participants possessed no known past or present injuries to their second or fourth fingers and no know hormonal abnormalities (according to parental report). The sample consisted of 62 males and 57 females with a mean age of 7.76 years (SD = 0.74) and 7.47 years (SD = 0.82) respectively. As reported in chapter 7 and 8 however, within this sample right hand 2D:4D data was missing for 1 female from experiment 5, and one female from experiment 6. Left hand 2D:4D data was also missing for one female from experiment 6.

9.2.3 Measures

9.2.3.1 SAT scores

See chapter 7 (section 7.2.6) for general information on the nature of KS1 SAT numeracy and literacy scores. KS1 SATs scores are rated in ascending order as 1, 2C, 2B, 2A, 3, 4 with 1 being the lowest and 4 the highest. As detailed by Brosnan (2008) SATs ratings are coded as 9, 13, 15, 17, 21, and 27 respectively, for the purpose of national data reporting. These coding were adopted in the current study and, similar to Brosnan (2008) numeracy and literacy scores were standardised into z-scores.

9.2.3.2 Second to Fourth Digit Ratio measurement

See chapters 7 (section 7.2.3) and 8 (section 8.2.3) for details regarding the reliability of 2D:4D measurement in the two data samples combined to form the current data set.

While males showed lower 2D:4D ratios (mean right hand 2D:4D = 0.954, SD = 0.03; mean left hand 2D:4D = 0.948, SD = 0.03) in comparison to females (mean right hand 2D:4D = 0.965, SD = 0.04; mean left hand 2D:4D = 0.957, SD = 0.03) sex differences in 2D:4D were not found to be significant for either the right ($t_{(115)} = 1.603$; $p = 0.112$) or left hand ($t_{(116)} = 1.624$; $p = 0.107$).

9.3 Results

In line with Brosnan (2008) analysis of SAT data was conducted on standardised Z-scores of the coded SAT scores. Means and standard deviations of coded SAT scores and standardised Z-coded SAT scores are displayed in table 60.

Table 60

Means and standard deviations (SD) of coded SAT numeracy and literacy scores and subsequent Z-SAT scores.

	Coded SAT scores		**Z-SAT scores**	
	numeracy	**literacy**	**numeracy**	**literacy**
Males (*n* = 62)	16.42 (3.27)	15.61 (3.74)	0.056 (0.95)	-0.072 (1.03)
Female (*n* = 57)	16.02 (3.63)	16.16 (3.52)	-0.061 (1.06)	0.078 (0.97)
Overall (*n* = 119)	16.23 (3.44)	15.87 (3.63)		

Kolmogorov-Smirnov analysis revealed that normality was violated for both Z-SAT numeracy and Z-SAT literacy scores. Mann-Whitney *U* tests therefore were used to explore sex differences in SAT performance and Spearman's correlations used to investigate any potential associations between SAT performance and age and SAT performance and 2D:4D.

In order to overcome the possibility that any discrepancies in the findings of the current study as compared to the results of Brosnan (2008) may arise due to different approaches to the analysis, analyses directly mirroring that of Brosnan (2008) were also conducted, a summary of this analysis can be found in Appendix 15.

A significant positive correlation was revealed between age and Z-SAT numeracy scores, $\rho = 0.193$; $p = 0.035$, suggesting higher scores with increasing age.

The correlation between age and Z-SAT literacy scores was not significant, $\rho = 0.014$, p = 0.88. No significant sex difference were identified in either Z-SAT numeracy, U = 1780, Z = -0.323, p = 0.747, or Z-SAT literacy, U = 1630, Z = -0.747, p = 0.455.

As can been seen in table 61 Spearman's correlation analysis revealed a significant correlation between right hand 2D:4D and Z-SAT literacy scores in females. The correlation was in a positive direction thus higher 2D:4D was associated with high Z-SAT literacy scores. No further significant correlation were revealed between 2D:4D and Z-SAT literacy scores. No significant associations were observed between 2D:4D and Z-SAT numeracy scores in analysis of the entire data set or male and female data analysed separately. As in previous sample sizes power for each Spearman's analysis was calculated in G*Power 3.1.2 (Faul et al., 2009) using adjusted sample sizes (see chapter 2, section 2.3.1, p. 51).

Table 61

Spearman's correlation coefficients (ρ), p values and power values for the relationship between left and right hand 2D:4D and Z-SAT numeracy and literacy scores. Significant effect highlighted in bold.

	Males and Females n = 119		**Males n = 62**		**Females n = 57**	
	Right 2D:4D	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***	***Right 2D:4D***	***Left 2D:4D***
Z-SAT Numeracy	$\rho = 0.084$ p = 0.366 1-β = 0.14	$\rho = 0.024$ p = 0.799 1-β = 0.057	$\rho = -0.009$ p = 0.944 1-β = 0.051	$\rho = 0.063$ p = 0.624 1-β = 0.075	$\rho = 0.150$ p = 0.274 1-β = 0.186	$\rho = -0.016$ p = 0.905 1-β = 0.051
Z-SAT Literacy	$\rho = 0.169$ p = 0.069 1-β = 0.422	$\rho = 0.062$ p = 0.504 1-β = 0.098	$\rho = -0.072$ p = 0.578 1-β = 0.083	$\rho = -0.006$ p = 0.961 1-β = 0.05	**$\rho = 0.343$** **p = 0.01** **1-β = 0.715**	$\rho = 0.112$ p = 0.369 1-β = 0.124

9.4 Discussion

The present experiment aimed to examine the possible relationship between 2D:4D and Key Stage 1 (KS1) SAT numeracy and literacy performance. The findings revealed a significant positive correlation between right hand 2D:4D and Z-SAT literacy scores in females suggesting the presence of higher literacy scores in higher 2D:4D participants. No significant correlations were revealed for either males or females for analysis for Z-SAT numeracy data.

Previous research by Brosnan (2008) reported a significant negative relationship between 2D:4D and: SAT numerical performance and numeracy performance relative to literacy performance in males and a significant positive association between 2D:4D and SAT literacy performance in females. No significant relationship between SAT literacy scores were identified in males and no significant relationship between either SAT numeracy performance or SAT numeracy relative to literacy were reported in females. The findings suggested a facilitative influence of PT on numerical skills in boys and a detrimental influence of PT on verbal competencies in girls. While the findings of the present study are partially in line with those describe by Brosnan (2008) in that a significant positive association was found between 2D:4D and SAT literacy scores in females the current experiment failed to replicate significant associations between 2D:4D and numeracy in males.

It is possible that the discrepancy between the findings of the current experiment and those reported by Brosnan (2008) may be a consequence of the two different adopted analyses thus data from the present experiment was re-analysed according to the procedure adopted by Brosnan (2008), see Appendix 15. The experiment conducted by Brosnan (2008) also investigated the relationship between right hand 2D:4D minus left hand 2D:4D (referred to as Dr-1) and SAT performance. As described in chapter 4 there is evidence that the sexual dimorphism in 2D:4D may be larger on the right hand relative to the left (Manning et al., 1998; Williams et al., 2000) there is also evidence that associations between 2D:4D and behavioural and cognitive outcomes are commonly reported to be stronger on the right hand. While the precise reason for this remains unknown, one explanation offered by Manning (2002) is that sexually dimorphic traits in general tend to be expressed more often in the “male form” on the right side of the body (Tanner, 1990). In line with this hypothesis there is evidence that paired organs may occasionally show directional asymmetry (Mittwoch & Mahadevaiah, 1980) and that the extent of this asymmetry may be associated with sex dependent cognitive patterns (Kimura, 1994). It has thus been hypothesised that Dr-1 may be negatively related to PT exposure. Brosnan (2008) found no significant correlations between Dr-1 and SAT performance, re-analysis of results from the current experiment however (according to the procedure adopted by Brosnan (2008)) found significant positive relationships between Dr-1 and; SAT numeracy and SAT literacy scores in females (see Appendix 15). Similar to findings with regard to right hand 2D:4D data the results suggest a facilitative impact of PT exposure on literacy in females. In addition, the findings for analysis of Dr-1 also imply a facilitative influence of PT on numeracy in females.

The lack of a significant relationship between measures of 2D:4D and SAT numeracy performance is also contrary to the findings of Fink et al. (2006) who reported a significant negative relationship between 2D:4D and basic numerical skills in boys aged 6-11. Evidence from the current experiment that PT may have a negative impact on numerical performance in females is also in opposition to the findings of Kempel et al. (2005) who reported an association between low 2D:4D (high PT) and improved performance on a numerical IQ task in females. The results however are in line with evidence presented by Finegan et al. (1992). These authors reported a negative relationship between amniotic fluid testosterone levels on counting and number fact task performance in girls aged 4.

Similar to the findings of experiment 4-6, despite being in the anticipated direction (males < females), evidence from the current experiment failed to replicate well established sex differences in 2D:4D. Also consistent with previous evidence from the current thesis effects relating to a potential influence of 2D:4D were displayed for one hand only (right hand in the present study).

In summary the current study, revealed a significant positive relationship between 2D:4D and SAT literacy in females suggesting a negative impact of PT on SAT literacy scores. Contrary to the findings of Brosnan (2008) however no significant associations were found between 2D:4D and SAT numeracy scores in males. To the extent that Dr-1 reflects exposure to PT, findings of the reanalysis according to the methods adopted by Brosnan (2008) also suggested a detrimental effect of PT on SAT numeracy scores in females.

Chapter 10

General Discussion

Evidence derived from animal studies (e.g. Gorski et al., 1980; Guillamon et al., 1988), from the effects of atypical hormone exposure (e.g. Hampson et al., 1998; Hier & Crowley, 1982; Hines et al., 2003; Imperato-McGinley et al., 1991; Resnick et al., 1986), from direct measures of prenatal testosterone (PT) in umbilical cord blood and amniotic fluid (e.g. Finegan et al., 1992; Grimshaw, Sitarenios, & Finegan, 1995; Jacklin et al., 1988) and from proposed somatic markers of foetal exposure to testosterone (e.g. Loehlin & McFadden, 2003; Manning et al., 1998) suggests that testosterone and its metabolites play an important role in the organisation of various sexually dimorphic brain regions and the subsequent activation and promotion of certain cognitive skills. Although there is long standing speculation as to a potential association between exposure to PT and subsequent mathematical ability relatively few studies have attempted to investigate the relationship between the two. Evidence from the limited range of research that has been conducted presents equivocal results (e.g. Brosnan, 2008; Finegan et al., 1992; Fink et al., 2006; Kempel et al., 2005; Luxen & Buunk, 2005).

Mounting evidence suggests that mathematical and numerical aptitude be rooted in an innate and possibly evolved sense of number reflected in a primitive ability to represent and manipulate numerical quantity (Dehaene et al, 2003). Utilising the second to fourth finger ratio (2D:4D) as a potential proxy for PT this current thesis sought to systematically explore the relationship between PT and basic and so-called ‘core’ numerical competencies in children and adults.

Experiment 1 aimed to replicate and extend previous evidence for a relationship between 2D:4D and basic numerical competencies in children (Fink et al., 2006). The data revealed a significant positive correlation between right hand 2D:4D and number comparison subtask performance in females aged 5-7 years. High 2D:4D therefore was associated with higher scores on the subtask suggesting a facilitate influence of prenatal testosterone.

Experiment 2 investigated the relationship between 2D:4D, subitizing and counting in adults. A significant three way interaction was observed between right hand 2D:4D, process (subitizing vs. counting) and visual field of stimulus presentation. Both high and low 2D:4D participants showed faster subitizing reaction times in the left visual field with the pattern of lateralization shown to be slightly more pronounced in high 2D:4D participants. With regard to counting reaction times however, while low right hand 2D:4D participants showed a left visual field advantage, a right visual field advantage was observed in high right hand 2D:4D participants.

Following a number of key methodological amendments, experiment 3 reconsidered a possible relationship between 2D:4D and subitizing in adults. A significant four-way interaction between right hand 2D:4D, task (subitizing vs. non-numerical control), sex and response hand was found. Post-hoc analysis of this interaction identified a significant three way interaction between 2D:4D, sex and response hand for analysis of subitizing data. While both low and high right hand 2D:4D males and females showed a right hand preference for subitizing, this preference was more pronounced in high right hand 2D:4D males relative to low right hand 2D:4D males, and low right hand 2D:4D females relative to high right hand 2D:4D females. Further analysis revealed a significant 2-way interaction between 2D:4D and response hand in females for subitizing reaction times. Closer inspection of this interaction showed a significant main effect of response hand in low 2D:4D females, no significant effect of response hand on subitizing reaction times was revealed for high 2D:4D females. Analysis of left hand 2D:4D data in experiment 3 also revealed a significant 3-way interaction between left hand 2D:4D, task (subitizing vs. control) and response hand. Both high and low left hand 2D:4D participants showed a right hand advantage on both the subitizing and control task. On the subitizing task the right hand advantage was slightly more pronounced in high 2D:4D participants, while on the control task the right hand advantage was more pronounced in low 2D:4D participants.

Taking into account further methodological limitations, experiment 4 re-examined the possible relationship between 2D:4D and subitizing. The experiment also considered possible associations between 2D:4D and counting and number comparison task performance. In addition, the experiment attempted to replicate evidence from Bull and Benson (2006) for a possible relationship between 2D:4D and the Spatial Numerical Association of Response Codes (SNARC) effect. Significant correlations in a negative direction were found between left hand 2D:4D and counting reaction times in females. The results also revealed a significant task (numerical vs. control) x left hand 2D:4D interaction on subitizing percentage error scores; task x right hand 2D:4D interaction on counting percentage error scores; task x left hand 2D:4D interaction on counting reaction times; and task x right hand 2D:4D interaction on number comparison reaction times. All of the above interactions (except that between task and right hand 2D:4D for counting percentage error scores), suggested faster reaction times/higher accuracy in high 2D:4D (low PT) participants relative to low 2D:4D participants on the numerical task and faster reaction times/higher accuracy in low 2D:4D participants relative to high 2D:4D participants on the control task. In direct opposition to these findings however, the interaction between task and right hand 2D:4D for counting percentage error scores suggested greater accuracy in low 2D:4D participants on the counting task and greater accuracy in high 2D:4D participants on the control task. The results of experiment 4 also revealed a significant 3-way interaction between left hand 2D:4D, task and response hand on subitizing reaction times. Low 2D:4D participants showed a left visual field advantage while high 2D:4D participants showed a right visual field advantage for subitizing. Visual field differences were more pronounced in low 2D:4D participants. Data from experiment 4 however failed to replicate evidence for a potential association between 2D:4D and the SNARC effect.

Experiment 5 examined possible relationships specifically between core numerical skills (subitizing and number comparison) and 2D:4D in children. Correlation analysis revealed significant associations between; left hand 2D:4D and subitizing reaction times to the left visual field in males (negative direction), right hand 2D:4D and subitizing reaction times to the right visual field in females (positive direction), and left hand 2D:4D and subitizing percentage error scores in females (negative direction). ANOVA analysis also revealed a three-way interaction between left hand 2D:4D, task (subitizing vs. control) and visual field on subitizing reaction times. Low 2D:4D participants showed a right visual field advantage while high 2D:4D participants showed a left visual field advantage. A significant interaction between left hand 2D:4D, task (number comparison vs. control) and visual field for analysis of number comparison percentage error scores was also observed. In accord with the pattern of results revealed for subitizing reaction times, low 2D:4D participants showed a right visual field advantage, while high 2D:4D participants showed a very slight left visual field advantage.

Experiment 6 investigated the relationship between 2D:4D and performance on a standardised assessment of basic numerical competencies (the Dyscalculia Screener; Butterworth, 2003) in children. The results showed no significant associations between either right or left hand 2D:4D and performance on the test battery.

Utilising a combination of data from participants in experiments 5 and 6, experiment 7 attempted to replicate evidence from Brosnan (2008) for an relationship between 2D:4D and Key Stage 1 (KS1) Standard Assessment Task (SAT) scores. The findings revealed a significant positive correlation between right hand 2D:4D and Z-SAT literacy scores in female with higher 2D:4D associated with higher literacy scores.

10.1 Prenatal testosterone and numerical competency

10.1.1 Direct associations between 2D:4D and numerical skill

In non-clinical populations, previous research in adults suggests a negative correlation between 2D:4D and numerical skills (Luxen & Bunnk, 2005) and 2D:4D and spatial representations of number (the so-called SNARC effect; Bull & Benson, 2006) in males and females. Kempel et al. (2005) also reported a significant negative association between performance on a continuing numerical series task in females, however these authors found no significant association between numerical performance and 2D:4D in males. To the extent that 2D:4D provides an accurate reflection of PT, the results suggest the higher PT may be associated with superior numerical task performance.

In children similar findings have been reported in males. Fink et al. (2006) found significant correlations between 2D:4D and number knowledge, counting and visual number representation in boys aged 6-11. Similarly, Brosnan (2008) reported a significant negative correlation between 2D:4D and KS1 SAT numeracy scores in boys aged 6-7. Both Fink et al. (2006) and Brosnan (2008) however reported no significant relationships between 2D:4D and numerical test scores in females. Contradictory results are described by Finegan et al. (1992) who identified a negative relationship between levels of testosterone in amniotic fluid and number fact task performance in girls aged 4, suggesting a detrimental impact of PT on numerical skill in females.

Since the current programme of research was completed, further evidence from Bull, Davidson and Nordmann (2009) has been published regarding a potential relationship between 2D:4D and basic numerical skills in children. Bull et al. (2009) investigated the association between 2D:4D and basic arithmetic, number sense (counting, number knowledge, pattern recognition and estimation) and visuo-spatial skills in male and female children aged five. Results revealed a significant negative correlation between 2D:4D and arithmetic task performance in boys, and a significant negative correlation between 2D:4D and number sense and visuo-spatial skill in girls, suggesting a positive influence of PT on performance.

Contrary to previous evidence, the results of the experiments conducted in the present thesis have found little support for a direct relationship between 2D:4D and basic numerical performance in either children or adults. Where findings have suggested a possible link (i.e. experiment 1, 4 and 5) there is little observable consistency in the nature and direction of the revealed relationships with previous literature on the topic.

The pattern of data revealed for number comparison performance in females in experiment 1 whereby higher 2D:4D (low PT) was associated with improved performance was similar to findings reported by Finegan et al. (1992). The results however were contrary to other previous studies reporting an association between low 2D:4D and improved numerical performance in females (Bull et al. 2009; Bull & Benson, 2006; Kempel et al. 2005; Luxen & Bunnk, 2005). Furthermore a similar significant association was not revealed for children aged 8-11 years.

Also in contrast to the majority of previous evidence on the topic the significant negative correlations between left hand 2D:4D and counting reaction times in females in experiment 4 implies an association between low PT and improved performance. Similarly the task x 2D:4D interaction effects revealed for subitizing, counting and number comparison performance in experiment 4 also predominantly suggest a disadvantageous effect of PT on performance. Post-hoc analysis of the interactions however failed to reveal a significant main effect of 2D:4D on reaction times/accuracy for any of the numerical tasks analysed independently, the findings therefore purely imply a potential relationship between 2D:4D and numerical task performance relative to control.

In line with the majority of previous evidence the significant positive correlation between right hand 2D:4D and right visual field subitizing reaction times in females in experiment 5 suggests a beneficial effect of PT on subitizing performance. This is contradicted however by the negative correlation found between left hand 2D:4D and subitizing percentage error scores for information presented to the right visual field in females. The negative correlation revealed between left hand 2D:4D and subitizing reaction times to the left visual field in males in experiment 5 (suggesting a detrimental effect of PT on reaction times) is also again contrary to previous evidence.

Some consistency however was revealed with regard to experiment 7 for the reanalysis of the data according to the procedure adopted by Brosnan (2008). Findings revealed a significant negative relationship between Dr-1 (right hand 2D:4D minus left hand 2D:4D) and SAT literacy and numeracy scores in females. As Dr-1 is thought to be negatively related to PT exposure the results potentially suggest a facilitative influence of PT on SAT numeracy scores. While the pattern of results however is similar to that identified by Brosnan (2008), contradictory to previous evidence, a similar relationship was not identified in males. It is also important to note that evidence from Brosnan (2008) does not report any associations between SAT performance and Dr-1. The significant associations reported by Brosnan (2008) relate purely to average 2D:4D (averaged from both the left and right hands). In the current study the association between 2D:4D (left, right or average) was not found to be significant. Thus while the implications of the findings relating to females in the current study are similar to those relating to males reported by Brosnan (2008) the findings are not directly comparable.

10.1.2 Prenatal testosterone and lateralization during basic numerical tasks

One finding which has been repeatedly identified in the current thesis is a possible association between 2D:4D and lateralization for core numerical processing. The most convincing evidence for a possible relationship between the two can be derived from the results of experiment 3. Based on the notion that hand preferences may reflect patterns of lateralization, the findings potentially suggest that high PT exposure in females may relate to increased lateralization during subitizing, and that the effects of PT on lateralization during subitizing in females may be different to those observed in males. This particular effect however was not replicated in subsequent experiments. Furthermore, experiment 3 also revealed a significant three way interaction between right hand 2D:4D, sex and response hand for analysis of control task reaction times. Thus, while the, pattern of results revealed for the control task was dissimilar to that observed for subitizing reaction times, it is important to be aware the a link between 2D:4D and lateralization may not be specific to numerical task performance.

The results of experiments 2-4 all revealed significant three way interactions between 2D:4D, task (numerical vs. control) and lateralization (as assessed by either visual field or response hand preferences) on subitizing task performance. Data from experiment 5 also revealed a significant interaction between 2D:4D, task and lateralization for both subitizing and number comparison task performance. Given that the basis of functional hemispheric asymmetry has long been hypothesised to be associated with exposure to PT (Hines & Shipley, 1984; Geschwind & Galaburda, 1987; Witelson, 1991) such findings are intriguing and potentially suggest a role for PT in patterns of functional neural organisation during basic numerical processing. However, while the findings relating to the relationship between 2D:4D and lateralization for subitizing and number comparison in experiment 5 showed similarities, generally evidence from experiments 2-5 demonstrates little consistency in the nature of the interactions between 2D:4D, visual field and task (subitizing vs. counting/control). Furthermore, closer inspection of all such interactions did not reveal a significant interaction effect of 2D:4D and visual field on either subitizing or control task performance analysed separately. The findings therefore purely suggest that 2D:4D differences in visual field preferences during core numerical processing are different to those demonstrated for similar reaction time tasks.

10.1.3 Sex differences in the nature of reported effects

It is also worth highlighting that there is ongoing inconsistency in previous evidence and the current study with regard to the nature of reported effects depending upon sex. Previous research from Fink et al. (2006), Brosnan (2008) and Kempel et al. (2005) have reported effects in one sex only while other researchers have reported effects in both males and females (e.g. Bull et al., 2009; Bull & Benson, 2006; Luxen & Bunnk, 2005). In the current thesis, the significant correlation between 2D:4D and number comparison in experiment 1 as well as the significant correlations between 2D:4D and counting task performance in experiment 4 and the potential influence of 2D:4D on lateralization for subitizing revealed in experiment 3 were all only found to be significant in females. In experiment 5 significant correlations in opposite direction between males and females were found for relationships between 2D:4D and subitizing reaction times. In females however a significant correlation for subitizing percentage error score was also found in a similar direction to the significant association revealed for male subitizing reaction times. As the issues of sex differences in the potential effects of PT on numerical intuition remains a source of debate it is important that the factor of sex is given appropriate consideration during analysis of future research on the topic.

10.1.4 Summary

In summary, throughout the thesis no clear pattern of results exists for an association between 2D:4D and basic numerical skill in general or between 2D:4D and any particular aspect of basic numerical competency (i.e. subitizing, counting, number comparison etc.). While it is important to recognise that a number of the experiments conducted in the thesis contained important methodological limitations, the ongoing lack of consistency in results is difficult to entirely explain with reference to potential methodological confounds. It is likely therefore that the erratic pattern of results demonstrated both across different experiments utilising similar numerical tasks, and across the various numerical skills assessed, is at least partially a reflection of the complex nature of any potential relationship between PT numerical skill. Furthermore, it is not the case that the experiments described failed to replicate evidence only on tasks adopted purely in the current thesis. Experiment 1 failed to replicate the results of Fink et al. (2006) utilising a similar numerical test battery. Experiment 4 failed to replicate previous evidence for an association between 2D:4D and the SNARC effect despite using a similar task as that employed by Bull and Benson (2006). Finally, experiment 7 failed to precisely replicate an association between 2D:4D KS1 SAT numeracy performance (Brosnan, 2008) despite the fact that exactly same method of assessment was considered. The lack of consistency both within the thesis and with previous literature highlights the importance of replication for existing and future findings.

10.2 Implications for general associations between prenatal testosterone and cognition.

The lack of consistency in findings in the current thesis may not be particularly surprising when considered in the context of evidence for the potential relationship between PT and alternative cognitive skills. There remains ongoing debate with regard to the extent to which PT is associated with spatial cognition. As described in chapter 1, clinical studies in individuals who have experienced atypical hormone exposure has suggested a positive relationship between testosterone and spatial ability for levels of testosterone exposure up to and within a ‘typical’ male range (Hampson et al., 1998; Puts et al., 2008; Resnick et al., 1986; Imperato-McGinley et al., 1991; Heir & Crowley, 1982). Evidence for altered spatial task performance in girls with Congenital Adrenal Hyperplasia (CAH; a condition in which the individual is exposed to above normal androgen levels both pre- and neonatally) however has not been consistently identified, with some authors reporting impaired performance in CAH females, or no apparent differences in performance (Baker & Erhaedt, 1974; Helleday et al., 1994; Hines et al., 2003; McGuire et al., 1975).

Evidence for a positive relationship between PT and spatial task performance has also been identified in animals (e.g. Dawson et al., 1975; Roof & Havens, 1992; Williams et al., 1990) and opposite sex twins (Cole-Harding et al., 1988). The evidence for a relationship between PT and spatial cognition based on direct measures of hormone exposure via umbilical cord blood and amniotic fluid are also inconclusive. While evidence from Jacklin et al. (1988) and Finegan et al. (1992) suggested a significant negative relationship between exposure to PT and spatial ability in females (both studies failed to identify significant associations in boys), evidence described by Grimshaw, Sitarenios, & Finegan (1995) found a positive association between PT and reaction time and speed of rotation during a mental rotations task in girls using a rotation strategy (indicated via the strength of relationship between reaction time and figure orientation); a trend in the opposite direction was indicated in boy. Evidence from 2D:4D is equally mixed. While the majority of research suggests a negative relationship between 2D:4D and spatial task performance in males (hence a positive relationship with PT) (e.g. Collaer et al., 2007; Manning & Taylor, 2001; McFadden & Schubel, 2003; Peters et al., 2007; Sanders et al., 2005) and no relationship between the two in females (Austin et al., 2002; Coolican & Peters, 2003; Hampson et al., 2008; Manning & Taylor, 2001; McFadden & Shubel, 2003; Poulin et al., 2004**;** Sanders et al., 2005; van Anders & Hampson, 2005), there is also evidence for a positive or no relationship in males (Austin et al., 2002; Burton et al., 2005; Coolican & Peters, 2003; Hampson et al., 2008; Kempel et al., 2005; Poulin et al., 2004; Putz et al., 2004), and both a negative and positive relationship in females (Burton et al., 2005; Collaer et al., 2007; Csathó et al., 2001; Kempel et al. 2005; Peters et al., 2007; Poulin et al. 2004; Putz et al., 2004).

Evidence for possible effects of PT on verbal and linguistic abilities is far less common. Where significant effects of prenatal testosterone on verbal and linguistic competency have been investigated, once again little consistency emerges, either within, or across, methodologies. A range of research from both clinical populations and use of 2D:4D as a proxy marker of PT suggests that there may be no effect of PT on verbal competencies (Austin et al., 2002; Baker & Erhardt, 1974; Hier & Crowley, 1982; Kempel et al., 2005; McGuire et al., 1975; Resnick et al., 1986; Sinforiani et al., 1994). There is also evidence however for a possible negative relationship between the two factors (Brosnan, 2008; Lutchmaya et al., 2002; Luxen & Buunk, 2005). Similar to research relating to spatial ability and numerical competencies, there is evidence to suggest that the relationship between PT may be different in males and females or only present in one sex (Brosnan, 2008; Burton et al., 2005). In line with such evidence, associations between Dr-1 and SAT literacy scores in experiment 7 were only significant in females.

Finally with reference to aspects of behavioural and functional lateralization, results are again mixed. While evidence from 2D:4D and females with CAH imply a positive association between PT and incidence of left handedness or reduced degree of right handedness (Fink et al., 2004; Manning, Trivers, et al. 2000; Manning & Peters, 2009; Nass et al., 1987; Nicholls et al., 2008; but see Resnick et al., 1987), evidence from amniotic fluid studies demonstrates a significant positive correlation between PT and degree of right handedness (Grimshaw, Bryden, & Finegan, 1995). Evidence however from 2D:4D, amniotic fluid and opposite sex twin studies also suggests a positive relationship between PT and degree of functional lateralization (Bourne & Gray, 2009; Cohen-Bendahan et al., 2004; Grimshaw et al., 1995b, but see Cohen- Bendahan, 2005).

Interestingly analysis of data from experiment 5 revealed a significant interaction effect between right hand 2D:4D and visual field of stimulus presentation in which low 2D:4D (high PT) participants demonstrated faster right visual field response times relative to left while high 2D:4D (low PT) participants showed faster left visual field response times relative to right for subitizing and subitizing control task reaction times. As the factors of right hand 2D:4D and visual field did not also interact with task, the results may imply evidence for different patterns of lateralization in low vs. high 2D:4D groups and thus a possible generic influence of PT on response times to such speeded response tasks. Experiment 5 also revealed a significant left hand 2D:4D x sex x visual field interaction on number comparison and number comparison control percentage error scores. Here high 2D:4D males and low 2D:4D females showed a right visual field advantage, while low 2D:4D males and high 2D:4D females showed a left visual field advantage. These findings potentially suggest a sex dependant generic effect of PT on functional lateralization for that particular type of speeded response task. Given however the lack of consistency across the two findings, and the fact that similar results were not identified in other experiments conclusions relating to these interactions remain extremely tenuous.

As previously discussed, evidence from the current thesis has offered very little support for an association between 2D:4D and performance on simple numerical tasks. For a number of the effects that have been however the relationship between 2D:4D and performance has often been identified only when considered relative to control. Such findings potentially highlight the possibility that PT may exert an influence upon various different cognitive tasks, the nature of which may differ depending upon the specific task under consideration. It is possible therefore that the lack of consistency across results relating to PT and cognition may be as least partially related to generic influence of PT on the various cognitive components that underlie performance on a particular task. If so, the necessity for control or consideration of the different processes that may determine task performance is absolutely essential in order to fully understand the potential influence of PT on a particular skill. Thus far, this is an issue that has been largely ignored by previous research, despite the fact that a wide range of different spatial, verbal, numerical and mathematical tasks have been adopted in order to assess any possible relationship between PT and cognition.

10.3 Sex differences in basic numerical processes

Previous evidence generally suggests a lack of sex differences in basic mathematical and numerical competencies (Brosnan, 2008; Bull et al., 2009; Fink et al., 2006; Geary, 1996). This finding however is not conclusive, evidence from Jordan et al. (2006) suggests a significant male advantage in kindergarten children on scores on a battery designed to test ‘number sense’. Males achieved higher overall scores on the test battery, and more specifically on the subsections designed to assess counting skills, number knowledge, non verbal calculation, estimation, and pattern recognition. Other studies have also presented evidence for a tendency for girls to use more language-based counting strategies in order to solve arithmetic problems, and a small advantage for boys on estimation tasks (e.g., Carr & Jessup, 1997; Jordan et al., 2003).

Experiment 2 demonstrated significant sex differences (female advantage) in counting task performance on the numerical test battery. Such sex differences however were only observed for children aged 8-11 years. As no further sex differences in performance were identified evidence from the present thesis is largely in line with previous evidence for a lack of sex differences in basic numerical skills.

10.4 Lateralization for basic numerical skills

There is evidence to suggest a certain degree of hemispheric specialization during core numerical processing. With regard to subitizing Pasini and Tessari (2001) reported a left visual field (LVF; right hemisphere) advantage for the identification and comparison of quantities in the subitizing range and a right visual field (RVF; left hemisphere) advantage for the comparison of quantities in the counting range. A right hemisphere advantage for subitizing has also been reported in other studies adopting behavioural measures of lateralization (Arp et al., 2006; Boles et al., 2007; Jackson & Coney, 2004). Contradictory evidence however was reported by Butterworth (1999) who found a left hemispheric advantage for subitizing on the basis of a single case study of an individual with brain damage. Evidence from brain imaging evidence also presents contrasting findings suggesting a lack of hemispheric specialization for the process (Piazza et al., 2002; Sathian et al., 1999).

In accord with previous behavioural evidence, the results of experiment 2 found a significant process x visual field interaction for the processes of subitizing vs. counting whereby a LVF advantage was revealed for the process of subitizing, and a RVF advantage was observed for the process of counting. Post-hoc analysis of this interaction however revealed that visual field differences were only significant for subitizing. This is in contrast to Pasini and Tessari (2001) who reported stronger visual field advantages for the process of counting. The method by which participants responded in experiment 2 however (all quantities in the subitizing range were responded to using the right hand, with all quantities in the counting range responded to with the left hand) is likely to have biased the revealed interaction. Furthermore, a right hemisphere advantage specifically relating to the process of subitizing was not replicated in any subsequent experiment. Generally therefore the results of the current thesis do not support evidence for right hemisphere lateralization during the process of subitizing.

While the results of experiment 4 failed to replicate evidence for a left visual field advantage during subitizing the experiment did identify a significant task x visual field interaction effect on counting task percentage error scores. Contrary to evidence from Pasini and Tessari (2001) however while a slight RVF advantage was evident on the control task, a LVF advantage was observed for performance on the counting task. Post-hoc analysis further revealed no significant effect of visual field on counting percentage error scores following Bonferroni correction.

With reference to our ability to approximately representation and compare numerical magnitude, there is evidence that parietal activation may be greater in the right hemisphere during a number comparison task (Chochon et al., 1999; Stanescu-Cosson et al., 2000). Experiment 4 also found a task x visual field interaction for reaction time on the number comparison task whereby, reaction times on the control task were faster for information presented to the RVF (left hemisphere) while average reaction times were faster for information presented to the LVF (right hemisphere) on the number comparison task. Such trends concur with evidence for greater right hemispheric activation during number comparison. Again however post-hoc analysis of the data showed no significant effect of visual field on number comparison reaction times following Bonferroni correction, and a similar association between task and visual field was not replicated in experiment 5.

10.5 Methodological and theoretical considerations

10.5.1 2D:4D as a marker of prenatal testosterone

Based on accumulating evidence, 2D:4D presents a potentially valuable tool to study the possible influence of prenatal sex hormones on subsequent behaviour and cognition. In contrast to studies exploring the cognitive correlates of atypical hormone exposure or hormones levels measured via umbilical cord blood or amniocentesis samples, research using 2D:4D is amenable to the use of controlled samples of any age, and provides a quick and simple-to-use measure that offers immediate data, with no ethical implications. As with all purported measures of PT however the technique is not without its limitations. As described in chapter 1, Putz et al. (2004) criticised the use of 2D:4D as a means by which to investigate the influence of PT on behavioural, cognitive and personality factors on the basis that there are numerous failures of replication in the many studies that have been conducted. In addition to this issue, there are a number of further criticisms regarding the use of 2D:4D which should be recognised, including: possible developmental changes in the ratio; problems isolating the specific biological mechanisms responsible for the trait; the paucity of direct evidence for 2D:4D as a marker of PT exposure; situations in which the sexual dimorphism for the trait is not identified; and inconsistencies relating to right vs. left hand results.

While sex differences in 2D:4D have been identified in children as young as two (Manning, 2002), the magnitude of these differences does appear to be smaller in children than adults (Manning et al, 1998, although see, McIntyre et al., 2005). A number of studies also imply a potential association between 2D:4D and age in children (Buck et al., 2003; Fink et al., 2004; McIntyre et al., 2005; 2006; Williams et al., 2003;) suggesting that 2D:4D may actually continue to change during periods of childhood growth. While such reported trends are typically weak or non-significant, more research is needed in order to clarify the possible impact of developmental changes on 2D:4D measurements. As potential age effects were not considered in the current thesis their possible influence on 2D:4D data in experiments adopting a young sample cannot be excluded.

The expression of 2D:4D may arise not entirely from prenatal androgens but as a consequence of alternative genetic, biological or environmental mechanisms. There is evidence for example to suggest a substantial genetic component to the expression of 2D:4D (Medland & Loehlin, 2008; Gobrogge, Breedlove & Klump, 2008; Paul, Kato, Cherkas, Andrew & Spector, 2006; Voracek & Dressler, 2007). While it is possible that any genetic determinants of 2D:4D may operate via a mechanism relating to the control of prenatal androgen exposure, it is equally possible that genes may influence expression of the trait by mechanisms entirely independent of exposure to androgens.

It is also important to recognise that evidence for an association between 2D:4D and PT accessed via routine amniocentesis may be more complex that is often recognised. As highlighted in chapter 1, Lutchmaya et al. (2004) found a significant negative correlation between the ratio of prenatal testosterone to estrogen from amniocentesis. While the research presents convincing support for the influence of fetal sex steroid levels on 2D:4D, the authors did not find significant associations between prenatal testosterone and prenatal estrogen. According to such evidence therefore the relationship between 2D:4D and psychological, behavioural and cognitive traits may actually reflect the impact of the balance between of fetal testosterone to estrogen as opposed to simply the level of PT exposure.

Evidence from Medland and Loehlin (2008) also identified possible environmental contributions to left and right hand digit ratio measures including poor placental or perinatal nutrition. Without comprehensive knowledge of the determinants of 2D:4D and the extent to which PT may regulate the trait beyond alternative genetic, hormonal and environmental factors the implications of a relationship between 2D:4D and cognitive, psychological and behavioural outcomes remain subject to speculation.

One further general issue regarding 2D:4D (although not specific to the measure) is the fact that, while support for the trait as an indication of PT is now fairly established existing evidence is based purely on data from alternative indirect measures of the exposure to the hormone. As evident in chapter 1 however there are various limitations associated with all indirect techniques that have been adopted in order to investigate effects relating to PT exposure. Unfortunately due to methodological restrictions, direct evidence for a relationship between fetal testosterone levels and 2D:4D remains elusive.

Of specific concern to the current thesis, 2D:4D is not entirely sexually dimorphic in that contrary classification of 2D:4D according to an individual's sex is not unusual (Beech and Beauvois, 2006). While a large body of evidence suggests that sex differences in the trait are relatively robust, Gobrogge et al. (2008) noted that such difference can be relatively subtle, with effect sizes ranging from 0.2-0.5. In the current thesis, experiments 1 and 4-7 all failed to replicate well established sex differences in 2D:4D. Failure to identify sex differences in 2D:4D is not unique to the current study (e.g. Austin et al., 2002; Brosnan, 2008; Bull et al., 2009), previous null findings however have typically been accounted for with reference to: population or ethnic confounds, reduced power as a result of low sample sizes or anomalous sampling. As experiments 4-7 controlled for the possible effects of within sample ethnic variation and sample sizes in experiment 1 (n = 73), 4-6 (n = 58-70) and, in particular experiment 7 (n = 119), were larger than those adopted in previous studies that have identified sex differences in 2D:4D (e.g. Kempel et al., 2005), ethnic variation and limited power do not offer convincing explanations for the failure to replicate sex differences in the current thesis. While it is possible that atypical sampling may present one potential explanation, it seems unusual that 4 consecutive experiments should have all recruited irregular participant samples. As both intra- and inter-rater reliability was revealed to be high in all experiments, measurement error is also unlikely to adequately account for the null results. Given that the critical period for the effect of hormones is thought to coincide with the male peak in testosterone you would expect sex differences during this period to be large. A consistent lack of sex differences therefore raises important questions as to the extent to which the measure reflects differences in PT in participants recruited in the current thesis.

Throughout the thesis there has also been a lack of consistency in results across measures of right and left hand 2D:4D. While a number of the revealed effects relate to analysis of the factor of left hand 2D:4D, others relate to the factor of right hand 2D:4D. Furthermore, none of the revealed significant effects relating to 2D:4D have been identified in analysis of both left and right hand 2D:4D measures. While there is evidence to suggest that the relationship between 2D:4D and psychological/cognitive factors may be stronger (or only present) for the right hand (e.g. Brown et al., 2002; Csathó et al., 2003; Lutchmaya et al., 2004; Williams et al., 2000), previous research does not provide any explanation for why revealed effect relating to left hand data may be absent for analysis of right hand data. Given that both measures are presumed to reflect PT exposure, it remains unclear why the two should demonstrate different associations with the same variables in identical analyses. Interestingly there is some evidence that heritability of 2D:4D may be greater for the left hand than the right (Gobrogge et al., 2008). Medland and Loehlin (2008) also reported significant non-shared genetic and environment variations between left and right hand 2D:4D measures. It is possible therefore the correlations between cognitive, behavioural and personality factors and 2D:4D may reflect different biological and environmental correlates depending upon whether the relationship is between right or left hand 2D:4D measures. Once again, in the absence of a more complete understanding of the determinants of 2D:4D concrete conclusions as to the nature of any revealed correlations is difficult, particularly when different effects are identified for right and left hand 2D:4D data.

10.5.2 Hormone effects

10.5.2.1 Organisational vs. Activational

As described in chapter 1, the actions of sex steroids are generally classified into organisational or activational effects. Organisational effects refer to those that occur during critical pre- or peri-natal periods to produce irreversible influences on the brain and behaviour, while activational effects refer to the acute, phasic influences of sex hormones as a result of circulating hormone levels. Research in animals has demonstrated that activational effects are often essential in order to allow the tissue or organ in question to perform its function (Pheonix et al. 1959, see chapter 1).

Despite being a useful heuristic it is now widely recognised that this classic dichotomy between organisational and activational effects is over-simplistic. In the past it was generally assumed that activational effects occur during adulthood, whereas organisational effects occur during fetal and immediate post-fetal development (Buchanan et al, 1992). An expanding body of literature however suggests that adolescence may actually constitute a second period of critical development during which sex hormones can exert a secondary organisational effect on the nervous system (e.g. Romeo, 2005; Sisk & Zehr, 2006). Such effects are typically characterised as building on or completing effects, as opposed to an entirely separate period of major organisation. While the majority of researchers however recognise the limitations of the dichotomy between activational and organisational effects, the potential interaction between the two is rarely taken into account in experimental design.

It is likely that a more complex model of potential hormone effects is necessary in order to comprehensively frame the impact of sex steroid hormones on neural, cognitive and behavioural function. Any contemporary model should look to incorporate all indentified gonadal hormone influences on the brain, physiology and cognition including, for example, 1) effects arising purely due to a prenatal organising effect, 2) organisational effects that also rely on elevated levels during puberty in order to be expressed, 3) effects arising purely due to a pubertal organising influence and, 4) effects arising purely due to circulating levels at a specific time point.

It is possible that the influence of a possible interaction between both pre-natal and pubertal hormones could account for some of the inconsistent findings in previous research. Intriguingly there is evidence that sex difference on certain spatial and mathematical tasks are more prevalent following adolescence (Hyde et al., 1990; Voyer et al., 1995). Although there are a range of social factors which are likely to contribute to this effect it is possible that biological factors that are associated with sex differences during this period, such as PT, may also play an important role in emerging sex difference in these competencies. While some authors are beginning to recognise and consider the potential importance of the joint contribution of both organisational and activational effects (e.g. Bourne & Gray, 2009) the area remains ripe for future investigation.

10.5.2.2 Association with estrogens and other hormones and neurotransmitters

As described above, one limitation of 2D:4D is the fact that the specific involvement of PT in development of the trait remains to be delineated from alternative biological mechanisms. One particular issue is the extent to which the trait is determined via fetal testosterone levels alone as opposed to the ratio of fetal testosterone to estrogen. While it is widely implicated that testosterone is responsible for the sexual differentiation of the fetus, there is increasing recognition of the importance of ovarian hormones for complete feminization. Evidence in animals suggests that that the removal of ovarian hormones may defeminze or masculinise neural development and subsequent cognition and behaviour; effects have been reported for open field activity, (Stewart & Cygan, 1980), mounting and lordosis (Dohler et al., 1984), the size of the corpus callosum (Fitch et al., 1991) and the SDN-POA (Dohler et al., 1984) and patterns of cortical thickness (Diamond et al., 1981). Possible relationships between 2D:4D and behavioural, cognitive, and psychological factors therefore could be at least partially related to variation in exposure to prenatal estrogens. Unfortunately however much remains unknown regarding the neural and cognitive effect of estrogen exposure during early development (Cohen-Bendahan et al., 2005).

It is also important to recognise that testosterone does not operate in isolation. Certain hormones which may also potentially affect cognition promote the secretion of testosterone, while testosterone itself may affect the secretion of other hormones (Cohen-Bendahan et al., 2005). In addition, steroid hormones may also have powerful effects on the synthesis and release of certain neurotransmitters (Neave, 2008). The possible effects of testosterone therefore may be far more complex and widespread that typically considered when discussing the potential relationships between PT and cognition.

10.5.3 Classifying core numerical competencies

At present there is no clear consensus on the precise competences which constitute core numerical skill. The precise conceptualisation and thus assessment of core numerical skill therefore can vary depending on a researcher's particular orientation. In the current thesis consideration of core numerical competencies was heavily based notion that origins of innate numerical knowledge are rooted in two systems of magnitude representation, 1) for the precise representation of small quantities (≤ 4; known as subitizing) and, 2) for the approximate representation of large quantities (Dehaene, 1997; Fiegenson et al., 2004; see chapter 2 for review).

Experiment 6 also considered associations between 2D:4D and an alternative, although similar, classification of core numerical competencies proposed by Butterworth (1999, 2005). According to Butterworth (1999) we are all born with an innate ‘number module’ characterized by the ability to: 1) understand that collections of things have a numerosity, that manipulations of sets can affect numerosity and that one collection has the same, greater or smaller numerosity as another, 2) understand that numerical collections need not be of visible things, and 3) recognise small numerosities (i.e. of collections up to about four objects) (see chapter 8).

Another related hypothesis as to the potential nature of innate numerical competencies has been described by Geary (1993; 1996). Geary (1993) refers to our innate capacity to possess fundamental quantitative competencies as “biologically primary quantitative abilities” which he categories into four broad domains of basic mathematics; 1) numerosity, the ability to quickly and accurately determine the quantity of small sets of items without the use of counting or estimating, 2) ordinality, a general sensitivity to more than and less than relations and later an understanding of specific ordinal relations, 3) counting, an implicit knowledge of the basic skeletal principle of one to one correspondence (i.e. that each item in an array can be tagged once and only once - Gelman and Gallistel 1978) and the pan-cultural understanding that serial-ordered number words can be used for counting, measurement and simple arithmetic and, 4) simple arithmetic, sensitivity to increases (addition) and decreases (subtraction) in the quantity of small sets.

While all three approaches share similarities there are certain differences in the range and precise nature of the innate numerical competencies which they identify. It is possible that by predominantly focusing on the ‘two core systems of numerical knowledge’ approach certain potentially innate numerical competencies which may present possible associations with PT may have been overlooked or inappropriately assessed.

10.5.4 Statistical analysis

One of most important limitations of the current study is that power appeared to be low across all of the experiments. While many of the non-significant effect sizes revealed for ANVOA analyses were extremely small, and thus arguably of limited practical application, a number of effect sizes relating to correlation analysis were within a small to medium range according to the conventions described by Cohen (1988). Such effect sizes are similar to those revealed in previous research exploring relations between 2D:4D and numerical performance (e.g. Brosnan, 2008; Fink et al., 2006).

Prospective power analysis suggests that in order to achieve a power of 0.8 for small to moderate effects using the two primary methods of analysis adopted in the current study, 196 parts are need for bivariate correlation and 260 for a 2 x 2 x 2 x 2 ANOVA analysis including two independent variables and two repeated variables, see Appendix 16 for the G*Power analysis relating to these calculations. These figures rise further when considering non-parametric analyses and, in the case of ANOVA, unbalanced designs (as seen in the current thesis). None of the experiments in the current thesis met the requirements for ANOVA. Similarly none of the experiments met the sample size requirements for correlation analysis once the sample was split by sex.

It should be noted however that where effect sizes were small to moderate or above the direction of revealed effects for the same task across different experiments and even for the same correlations across right and left hand measures were not always consistent. Furthermore in an attempt to address the issues relating to power in the current thesis subitizing reaction time data from experiments 2-4 and counting reaction time data from experiments 2 and 4 were combined, and associations between 2D:4D and performance re-investigated using the elevated sample sizes. As described in experiment 4, no significant correlations were found and the effect sizes observed following these reanalyses were actually lower than those observed for the same effects in analysis of the data separately within each experiment. Furthermore, it is worth highlighting that the sample sizes included in both the reanalysis of subitizing and counting data and in experiment 7 for the consideration of SAT numeracy scores were higher (in some cases by more than double) than those included in previous evidence where significant associations between 2D:4D and number performance have been identified (e.g. Bull et al., 2009; Brosnan, 2008; Fink et al. 2006).

While the issue of power may have reduced chances of findings a significant effect (given that the is a real effect to find) an limitation presenting the opposite problem is that throughout the thesis multiple correlations and ANOVAs have been employed in order to investigate potential 2D:4D effects on numerical performance. While corrections for multiple comparisons were applied to certain post-hoc analyses, such corrections were not adopted to control for multiple ANOVA analysis within each experiment. In certain experiments a particularly large number of analyses were required in order to consider the effects of both right and left hand digit ratios on the various different numerical tasks assessed. Controls for multiple comparisons in these experiments therefore would have been exceptionally stringent. As aspects of the research remain exploratory, it was felt that such controls may mask findings of potential interest in such early stages of understanding. The large number of analyses carried out in certain experiments however may have resulted in spurious findings as a consequence of inflated Type I error rates.

10.5.5 Extraneous variables

While it is impossible to control for all factors which could have potentially influenced numerical performance in the current thesis besides PT, there are obvious extraneous variable which should be considered. Firstly, as highlighted in chapter 1, there is some evidence that circulating levels of testosterone may be associated with performance on certain spatial (Christiansen, 1993; Christiansen & Knussmann, 1987; Gouchie & Kimura, 1991; Gordon & Lee, 1986; Moffat & Hampson, 1996; Neave et al., 1999; Shute, 1983), verbal (Christiansen, 1993; Christiansen & Knussmann, 1987), mathematical (Gouchie & Kimura, 1991) and simple reaction time (Müller, 1994) tasks. Such findings however are not consistently reported, and there are inconsistencies with regard to the revealed direction of the relationship with spatial task performance. While general consensus suggests that no significant relationships exist between 2D:4D and circulating testosterone measured via saliva (Bang et al., 2005; Hönekopp et al., 2007; Kallai et al., 2005; Kempel et al., 2005; Manning et al., 2004; Neave et al., 2003; van Anders & Hampson, 2005) evidence for a potential association between the two has also been identified by Manning et al. (1998). Any association between 2D:4D and circulating testosterone and/or circulating testosterone and performance could have possibly confounded the results. There is also evidence that performance on certain cognitive tasks may fluctuate with the phases of the menstrual cycle and oral contraceptive use in females (e.g. Hampson & Kimura, 1988; Hausmann, Slabbekoorn, van Goozen, Cohen-Kettenis & Güntürkün, 2001; Silverman & Phillips, 1993;). As such factors are associated with fluctuating hormone levels it is possible that failure to control for possible oral contraceptive and menstrual cycle effects may also have potentially influenced the association between 2D:4D and aspects of cognitive performance.

While a discussion of the potential social factors that may influence numerical performance are beyond the scope of the current thesis, it is important to recognise again that the possible influence of such factors is likely to have had an impact on performance in the current thesis. As highlighted in chapter 6 there is evidence to suggest that even core numerical skills may show an element of developmental progression (Svenson & Sjoberg, 1983), thus even in children the demonstration of such competencies is likely to represent an interaction between biological and sociocultural influences. It is also important to recognise that underlying biological dispositions may influence the environment which an individual is subject to (Geary, 1996). Individual variation in exposure to PT therefore may actually bias sociocultural influences. The interplay between biological and social factors that may impact upon cognition is thus incredibly complex and extremely difficult to break down.

10.6 Future research

As previously discussed, each of the different methodologies employed in order to assess potential associations between exposure to PT and cognition makes certain assumptions and has its own particular limitations. In order to reliability understand and interpret existing evidence for a relationship between PT and any aspect of cognitive performance, a convergence of evidence across different methodologies is required. With regard to core and basic numerical skills, research is predominantly based on evidence from 2D:4D, further research adopting a range of different techniques in order to test the possible relationship between the two factors therefore would be beneficial. As the current thesis however has consistently failed to identify convincing evidence for a potential relationship between 2D:4D and basic and core numerical task performance, it may be more fruitful to focus on the possible relationship between PT and higher mathematical tasks. Given that hormone exposure is known to have a profound effect on physical sexual dimorphisms and the majority of neural sex differences studied in animals (Arnold, 1996), tasks which show robust sex differences are likely to be the most promising candidates for a relationship with PT.

As the current thesis has failed to identify consistent evidence across different numerical tasks it is important that future research on the topic recognises that any revealed relationships between markers of PT and performance may be task specific and seeks to replicate any significant results. Evidence from the current thesis has also highlighted possible generic relationships between 2D:4D and reaction times, it is important then that future research also considers and, where relevant, controls for underlying cognitive processes that may mediate potential associations between PT and performance on the particular cognitive task of interest. With regard to mathematics, mathematical ability even in very young children is known to be related to spatial skills, verbal abilities, and working memory (see chapter 1), although difficult, it would be interesting to attempt to break down the different cognitive components involved in performance on different numerical tasks in order to consider potential relationships between PT and the various cognitive functions involved in performance.

As described in chapter 1, the most robust evidence for sex differences in the domain of mathematics can be derived from research in samples of mathematically gifted individuals. It would also be useful therefore to consider possible relationships between a marker of PT and numerical competencies in a mathematically gifted sample. At the opposite end of the spectrum, it would also be interesting to investigate any relationship between fetal testosterone levels and performance in individuals with mathematical learning difficulties.

Although inconsistent in their nature, potential associations between 2D:4D and lateralization for basic numerical performance relative to control have been identified in several studies through the current thesis. One possible interpretation of these findings is that PT may work to organise neural function or structure but that differences in activation may not necessarily translate to differences in performance. Evidence for such an effect would be in line with research showing that men and women may engage different constellations of brain regions to achieve the same level of performance on at least some cognitive and intelligence measures (Haier et al., 2005). It may be informative therefore to consider possible relationships between PT and functional brain activation during basic mathematical performance utilising brain imaging techniques.

Finally, as identified above, given that sex difference on certain mathematical tasks are more pronounce following approximately 13-16 years of age (Hyde et al.,

1990) it would be extremely interesting to explore the possible impact of pubertal hormone levels and the interaction between prenatal and pubertal hormone levels on mathematical tasks in which sex differences emerge during this time.

10.7 Conclusion

The current thesis has attempted to systematically explore potential relationships between 2D:4D as a somatic marker of PT exposure and basic numerical competencies in children and adults. While each experiment found some associations between 2D:4D and at least one aspect of numerical performance or lateralization for numerical performance, no observable pattern can be identified across the revealed effects. To the extent that 2D:4D reflects exposure to PT therefore, these findings suggest that any impact that PT may have on ‘core’ and basic numerical processing is likely to be complex and, given the lack of consensus in observed effects across the adopted tasks, task specific. Prior evidence has used a range of widely different numerical and mathematical assessments in order to assess relationships between correlates of PT and numerical and mathematical performance. In light of evidence presented in the present thesis, re-assessment and evaluation is vital before any conclusions and interpretations are drawn with regard to a link between PT and numerical or mathematical ability based on such evidence. Given the null findings in the current thesis, it may be more fruitful in future to focus on potential relationships in higher mathematical skills. Future research should recognise and attempt to consider the complexity of any possible associations between PT and numerical competencies by taking into account the underlying cognitive components which may determine task performance.

References

Aggleton, J. P., Kentridge, R. W., & Good, J. M. N. (1994). Handedness and musical ability: A study of professional orchestral players, composers and choir members. *Psychology of Music, 22,* 148–156.

Allen, J. S., Damasio, H., Grabowski, T. J., Bruss, J., & Zhang, W. (2003). Sexual dimorphism and asymmetries in the gray-white composition of the human cerebrum. *Neuroimage, 18,* 880-894.

Allen, L. S., Richey, M. F., Chai, Y. M., & Gorski, R. A. (1991). Sex differences in the corpus callosum of the living human being. *Journal of Neuroscience, 11,* 933-942.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders, DSM IV.* American Psychiatric Association: Washington, DC

Annett, M. (1985). Left, right, hand and brain: The right shift theory. Hillsdale, NJ: Lawrence Erlbaum Associates.

Annett, M. (1998). Handedness and cerebral dominance: the right shift theory. *Journal of Neuropsychiatry and Clinical Neurosciences, 10,* 459-469.

Annett, M. (1999). Left-handedness as a function of sex, maternal versus paternal inheritance, and report bias. *Behavior Genetics, 29,* 103-114.

Antell, S. E., & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development, 54,* 695-701.

Arai, Y., Nishizuko, M., Murakani, S., Miyakama, M., Machida, M., Takeuchi, H. & Sumida, H. (1993). Morphological correlates of neuronal plasticity to gonadal steroids: sexual differentiation of the preoptic area. In M. Hang, R. E. Whalen, C. Aron, & K. L. Olsen (Eds.) *The development of sex differences and similarities in behaviour* (pp. 311-323). Amsterdam: Kluwer Academic.

Arnold, A. P. (1996). Genetically triggered sexual differentiation of brain and behaviour. *Hormones and Behavior, 30,* 495-505.

Arp, S., Taranne, P, & Fagard, J. (2006). Global perception of small numeroisties (subitizing) in cerebral-palsied children. *Journal of Clinical and Experimental Neuropsychology, 28,* 405-419.

Austin, E. J., Manning, J. T., McInroy, K., & Mathews, E. (2002). A preliminary investigation of the association between personality, cognitive ability and digit ratio. *Personality and Individual Differences, 33,* 1115-1124.

Bachevalier, J. & Hagger, C. (1991). Sex differences in the development of learning abilities in primates. *Psychoneuroendocrinology, 16,* 177-188.

Badin, N. A. (1983). Arithmetic and non verbal learning. In: H. R. Myklebust (Ed) *Progress in learning disabilities, Volume 5.* New York: Grune and Stratton

Baker, S. W. & Ehrhardt, A. A. (1974). Prenatal androgen, intelligence and cognitive sex differences. In R. C. Friedman, R. M. Richart & R. L. Vande Wiele (Eds.) *Sex differences in behaviour* (pp. 53-76). New York: Wiley.

Balakrishnan, J. D., & Ashby, F. G. (1991). Is subitizing a unique numerical ability? *Perception and Psychophysics, 50,* 555-564.

Balakrishnan, J. D., & Ashby, F. G., (1992). Subitizing: Magical numbers or mere superstition? *Psychological Research, 54,* 80-90.

Bang, A. K., Carlsen, E., Holm, M., Petersen, J. H., Skakkebaek, N. E., & Jørgensen, N. (2005). A study of finger lengths, semen quality and sex hormones in 360 young men from the general Danish population. *Human Reproduction, 20,* 3109-3113.

Baron-Cohen, S., Luchmaya, S. & Knickmeyer, R. (2004). *Prenatal testosterone in mind: Amniotic fluid studies.* Massachusetts, U.S.A.: MIT Press.

Barrett, R. J. & Ray, O. S. (1970). Behaviour in the open field. Lashley III maze, shuttle box, and Sidman avoidance as a function of strain, sex and age. *Developmental Psychology, 3,* 73-77.

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representation in adults. *Cognition, 86,* 201-221.

Barth, H., Le Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition, 98,* 199-222.

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences, 102,* 14116-14121.

Beatty, W. W. (1979). Gonadal hormones and sex differences in non-reproductive behaviours in rodents: Organizational and activational influences. *Hormones and Behaviour*, *12*, 112-163.

Beatty, W. W. (1992). Gonadal hormones and sex differences in non-reproductive behaviours. In A. A. Gerall, H. Moltz & I. L. Ward (Eds.) *Handbook of neurobiology* (pp. 85-128). New York: Plenum Press.

Beech, J., & Beauvois, M. W. (2006). Early experience of sex hormones as a predictor of reading, phonology, and auditory perception. *Brain and Language, 96,* 49-58.

Benbow, C. P. (1986). Physiological correlates of extreme intellectual precocity. *Neuropsychologia, 24,* 719–725.

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral and Brain Sciences, 11,* 169-183.

Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or Artefact? *Science, 210,* 1262-1264.

Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science, 222,* 1029-1031.

Benoit, L., Lehalle, H., & Jouen, F. (2004). Do young children acquire number words through subitizing or counting? *Cognitive Development, 19,* 291-307.

Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology, 41,* 174-181.

Berenbaum, S. A. & Hines, M. (1992). Early androgens are related to childhood sex-typed toy preferences. *Psychological Science, 3,* 203-206.

Berenbaum, S. A., Moffat, S., Wisniewski, A. & Resnick, S. (2003). Neuroendocrinology: Cognitive effects of sex hormones. In M. De Haan & M. H. Johnson (Eds.) *The cognitive neuroscience of development* (pp. 207-235). New York: Psychology Press.

Bidlingmaier, F., Strom, T. M., Dorr, H. G., Eisenmenger, W. & Knorr, D. (1987). Estrone and estradiol concentrations in human ovaries, testes, and androgens during the first two years of life. *Journal of Clinical Endocrinology & Metabolism, 65,* 862-867.

Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do four-day-old infants categorize multisyllabic utterances. *Developmental Psychology, 29,* 711-721.

Bilger, R., Matthies, M. L., Hammel, D. R. & Demorest, M. E. (1990). Genetic implications of gender differences in the prevalence of spontaneous otoacoustic emissions. *Journal of Speech, Language and Hearing Research, 33,* 418-432.

Bishop, D. V. M. (1990). *Handedness and developmental disorder. Clinics in developmental medicine. Volume 110.* Philadelphia: J.B. Lippincott.

Bleeker, M. M., & Jacobs, J. E. (2004). Achievement in math and science: Do mothers' beliefs matter 12 years later? *Journal of Educational Psychology, 96,* 97-109.

Bleier, R., Byne, W. & Siggelkow, I. (1982). Citoarchitectonic sexual dimorphism of the medial preoptic area and anterior hypothalamic areas in the guinea pig, rat, hamster and mouse. *Journal of Comparative Neuropsychology, 212,* 118-130.

Boles, D. B., Phillips, J. B., & Givens, S. M. (2007). What dot clusters and bar graphs reveal: Subitizing is fast counting and subtraction. *Perception and Psychophysics, 69,* 913-922.

Boles, D. (2005). A large-sample study of sex differences in functional cerebral lateralization. *Journal of Clinical and Experimental Neuropsychology, 27,* 759-768.

Bourne, V., & Gray, D. L. (2009). Hormone exposure and functional lateralisation: Examining the contributions of prenatal and later life hormonal exposure. *Psychoneuroendocrinology, 34*, 1214-1221.

Brannon, E. M. (2006). The representation of numerical magnitude. *Current Opinion in Neurobiology*, *16,* 222-229.

Brannon, E. M., Abbott, S., & Lutz, D. J. (2004). Number bias for the discrimination of large visual sets in infancy. *Cognition, 93,* B59-B68.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282,* 746-747.

Breedlove, S. M. (1992). Sexual dimorphism in the vertebrate nervous system. *Journal of Neuroscience, 12*, 4133-4142.

Brosnan, M. J. (2008). Digit ratio as an indicator of numeracy relative to literacy in 7-year-old
British schoolchildren. *British Journal of Psychology, 99*, 75-85.

Brown, W.M., Finn, C. & Breedlove, S.M. (2002). Sexual dimorphism in digit-length ratios of laboratory mice. *Anatomical Record, 267*, 231–234.

Brown, W. M., Hines, M., Fane, B. A., & Breedlove, S. M. (2002). Masculinised finger length patterns in human males and females with congenital adrenal hyperplasia. *Hormones and Behavior, 42,* 380-386.

Bryden, M. P., McManus, I. C., & Bulman-Fleming, M. B. (1994). Evaluating the empirical support for the Geschwind–Behan–Galaburda model of cerebral lateralization. *Brain and Cognition, 26*, 103–167.

Buchanan, C. M., Eccles, J. S., Becker, J. B. (1992) Are adolescents victims of raging hormones? Evidence for activational effects of hormones on moods and behaviour at adolescence. *Psychological Bulletin, 111,* 62-107.

Buchsbaum, M. S. & Henkin, R. I. (1980). Perceptual abnormalities in patients with chromatin negative gonadal dysgenesis and hypogonadatropic hypogonadism. *International Journal of Neuroscience, 11,* 201-209.

Buck, J. J., Williams, R. M., Hughes, I. A., & Acerini, C. L. (2003). In-utero androgen exposure and 2nd to 4th digit length ratio - Comparisons between healthy controls and females with classical congenital adrenal hyperplasia. *Human Reproduction, 18,* 976–979.

Bull, R., & Benson, P. J. (2006). Digit ratio (2D:4D) and the spatial representation of magnitude. *Hormones and Behavior, 50,* 194-199.

Bull, R., Davidson, W. A., & Nordmann, E. (2009). Prenatal testosterone, visual-spatial memory, and numerical skills in young children. *Learning and Individual Differences,* dio:10.1016/j.lindif.2009.12.002

Bull, R., & Johnston, R. S. (1997). Children’s arithmetical difficulties: Contributions from processing speed, item identification, and short- term memory. *Journal of Experimental Child Psychology, 65,* 1-24.

Bull, R., Johnston, R. S. & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch and central executive in children’s arithmetical skills: Views from cognition and experimental neuropsychology. *Developmental Neuropsychology, 15,* 421-442.

Bull, R., Sherif, G. (2001). Executive functioning as a predictor of children’s mathematical ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*, 273-293.

Burbaud, P., Camus, O., Guehl, D., Bioulac, B., Caille, J. M., & Allard, M. (1999). A functional magnetic resonance imaging study of mental subtraction in human subjects. *Neuroscience Letters, 273,* 195-199.

Burns, E. M., Campbell, S. L., Arehart, K. H. & Keefe, D. H. (1993). Long-term stability of spontaneous otoacoustic emissions (Abstract). *Association for Research in Otolaryngology, 16,* 98.

Burns, E. M., Campbell, S. L. & Arehart, K. H. (1994) Longitudinal measurements of spontaneous otoacoustic emissions in infants. *Journal of the Acoustic Society of America, 95,* 385-394.

Burton, L. A., Henninger, D. & Hafetz, J. (2005). Gender differences in relations of mental rotation, verbal fluency, and SAT scores to finger length ratios as hormonal indexes. *Developmental Neuropsychology, 28,* 493-505.

Butterworth, B. (1999). *The Mathematical Brain*. London: Macmillan.

Butterworth, B., Cappelletti, M., & Kopelman, M. (2001). Category specificity in reading and writing: The case of number words. *Nature Neuroscience, 4,* 784-786.

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science, 16,* 572-578.

Calnan, M., & Richardson, K. (1976). Developmental correlates of handedness in a national sample of 11-year-ilds. *Annals of Human Biology, 3,* 329-342.

Cantlon, J. F., & Brannon, E. M. (2007) Basic maths in monkeys and college students. *PLoS Biology, 5,* e328. dio: 10. 1371/journal.pbio.0050378

Cappa, S. F., Guariglia, C., Papagno, C., Pizzamiglio, L., Vallar, G., Zoccolotti, P., Ambrosi, B. & Santiemma, V. (1988). Patterns of lateralization and performance levels for verbal and spatial tasks in congenital androgen deficiency. *Behavioural Brain Research, 31,* 177-183.

Cappelletti, M., Butterworth, B., & Kopelman, M. (2001). Spared numerical abilities in a case of semantic dementia. *Neuropsychologia, 39,* 1224-39.

Carr, M., & Jessup, D. L. (1997). Gender differences in first-grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology, 89,* 318–328.

Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology, 31*, 697-705.

Chada, M., Prusa, R., Bronsky, J., Pechova, M., Kptaska, K. & Lisa, L. (2003). Inhibin B, follicle stimulating hormone, luteinizing hormone, and estradiol and their relationship to the regulation of follicle development in girls during childhood and puberty. *Physiological Research, 52,* 341-346.

Chi, M. T. C., & Klahr, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology, 29,* 434-439.

Chiang, W. C., & Wynn, K. (2000). Infant tracking of objects and collections. *Cognition, 77,* 169-195.

Chochon, F., Cohen, L., van der Moortele, P. F., & Dehaene, S. (1999) Differential contributions of the left and right inferior parietal lobules to number processing. *Journal of Cognitive Neuroscience, 11,* 617-630.

Choi, J. & Silverman, I. (1997). Sex dimorphism in spatial behaviours applications to route learning. *Evolution and Cognition, 2,* 165–171.

Christiansen, K. (1993). Sex hormone-related variations of cognitive performance in !Kung San hunter-gatherers of Namibia. *Neuropsychobiology, 27,* 97-107.

Christiansen, K., & Knussmann, R. (1987). Sex hormones and cognitive functioning in men. *Neuropsychobiology, 18,* 27-36

Church, R. M., & Meck, W. H. (1984). The numerical attribute of stimuli. In: H. L. Roitblatt, T. G. Bever, H. S. Terrace (Eds.) *Animal cognition.* Erlbaum: Hillsdale, NJ

Cipolotti, L., Butterworth, B., & Denes, G. (1991). A specific deficit for numbers in a case of dense acalculia. *Brain, 114,* 2619–2637.

Clark, J. H., Schrader, W. T. & O'Mally, B. W. (1985). Mechanisms of steroid hormone action. In J. D. Wilson & D. W. Foster (Eds.) *Williams textbook of endocrinology* (pp. 33-75). Philadelphia, U.S.A.: W. B. Saunders.

Clark, M. M. & Galef, B. G. (1998). Effects of intrauterine position on the behaviour and genital morphology of litter-bearing rodents. *Developmental Neuropsychology, 14,* 197-211.

Clark-Carter, D. (2010). *Quantitative Psychological Research. Third Edition.* East Sussex: Psychology Press.

Clearfield, M. W., & Mix, K. S. (1999). Number versus Contour Length in Infants' Discrimination of Small Visual Sets. *Psychological Science, 10,* 408-411.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences (2nd edition).* Hillsdale, NJ: Erlbaum

Cohen-Bendahan, C. C. C., Buitelaar, J. K., Van Goozen, S. H. M, & Cohen-Kettenis, P. T. (2004). Prenatal exposure to testosterone and functional cerebral lateralization: A study in same-sex and opposite-sex twin girls. *Psychoneuroendocrinology, 29,* 911-916.

Cohen-Bendahan, C. C. C., van de Beek, C. & Berenbaum, S. A. (2005). Prenatal sex hormone effects on child and adult sex-typed behaviour: methods and findings. *Neuroscience and Biobehavioral Reviews, 29,* 353-384.

Cohen-Bendahan, C. C. C. (2005). Biological roots of sex differences: A longitudinal twin study. University Utrecht: Ph.D. thesis. In Cohen-Bendahan, C. C. C., van de Beek, C. & Berenbaum, S. A. (2005). Prenatal sex hormone effects on child and adult sex-typed behaviour: methods and findings. *Neuroscience and Biobehavioral Reviews, 29,* 353-384.

Cole-Harding, S., Morstad, A. L. & Wilson, J. R. (1988). Spatial ability in members of opposite-sex twin pairs (Abstract). *Behaviour Genetics, 18,* 710.

Collaer, M. L., & Hines, M. (1995). Human behavioural sex differences: A role for gonadal hormones during early development. *Psychological Bulletin*, *118*, 55-107.

Collaer, M. L., Reimers, S. & Manning, J. T. (2007). Visuospatial performance on an internet line judgment task and potential hormonal markers: Sex, sexual orientation, and 2D:4D. *Archives of Sexual Behaviour, 36*, 177-192.

Collins, D. W. & Kimura, D. (1997). A large sex difference on a two-dimensional mental rotation task. *Behavioral Neuroscience, 111,* 845–849.

Colvin, M. K., Funnell, M. G., & Gazzaniga, M. S. (2005). Numerical processing in the two hemispheres: Studies of a split-brain patient. *Brain and Cognition, 57,* 43-52.

Coolican, J., & Peters, M. (2003). Sexual dimorphism in the 2D/4D ratio and its relation to mental rotation performance. *Evolution and Human Behavior, 24*, 179-183.

Crow, T., Crow, L., Done, D., & Leask, S. (1998). Relative hand skill predicts academic ability: global deficits at the point of hemispheric indecision. *Neuropsychologia, 36,* 1275–1282

Csathó, A., Osváth, A., Bicsák, E., Karádi, K., Manning, J., & Kállai, J. (2003). Sex role identity related to the ratio of second to fourth digit length in women. *Biological Psychology, 62,* 147-156.

Csathó, A., Osváth, A., Karádi, K., Bicsák, E., Manning, J., & Kállai, J. (2001). Spatial navigation related to the second to fourth digit length in women. *Learning and Individual Differences, 13,* 239-249.

Dane, S., & Balci, N. (2007). Handedness, eyeness and nasal cycle in children with autism. *International Journal of Developmental Neuroscience, 25,* 223-226.

Danilovic, D. L. S., Carrea, P. H. S., Casto, E. M. F., Melo, K. F. S., Mendonca, B. B. & Arnhold, I. J. P. (2007). Height and bone mineral density in androgen insensitivity syndrome with mutations in the androgen receptor gene. *Osteoporosis International, 18,* 369-374.

Dassonville, P., Zhu, X-H., Uğurbil, K., Kim, S-G., & Ashe, J. (1997). Functional activation in motor cortex reflects the direction and degree of handedness. *Proceedings of the National Academy of Sciences of the United States of America, 94*, 14015-14018.

Davatzikos, C., & Resnick, S. M. (1998). Sex differences in anatomic measures of interhemispheric connectivity: Correlations with cognition in women but not men. *Cerebral Cortex, 8*, 635-640.

Dawood, M. Y. & Saxena, B. B. (1977). Testosterone and dihydrotestosterone in maternal and cord blood and in amniotic fluid. *American Journal of Obstetrics and Gynaecology, 129,* 37-42.

Dawson, J. L. M., Cheung, Y. M. & Lau, R. T. S. (1975). Developmental effects of neonatal sex hormones on spatial and activity skills in the white rat. *Biological Psychology, 3,* 213-229.

Dehaene, S. (1996). The organization of brain activations in number comparison: Event-related potential and the additive-factors method. *Journal of Cognitive Neuroscience, 8*, 47-68.

Dehaene, S. (1997). *The Number Sense: How the mind creates mathematics.* New York: Oxford University Press

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371-396.

Dehaene, S., & Cohen, L. (1994). Dissociable mechanisms of subitizing and counting: Neuropsychological evidence from simultanagnostic patients. *Journal of Experimental Psychology: General, 20,* 958-975.

Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex, 33,* 219-250.

Dehaene, S. (2000). Cerebral bases of number processing and calculation. In M. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (pp.987-998). Cambridge, MA: MIT Press.

Dehaene, S., & Changeux, J-P. (1993). Development of Elementary Numerical Abilities: A Neuronal Model. *Journal of Cognitive Neuroscience, 5,* 390-407.

Dehaene, S., & Marques, F. (2002). Cognitive euroscience: Scalar variability in price estimation and the cognitive consequences of switching to the euro. *Quarterly Journal of Experimental Psychology, 55A,* 705-731.

Dehaene, S., Piazza, M. Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487-506

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioural and brain-imaging evidence. *Science, 284,* 970-974

Dehaene, S., Tzourio, N., Frak, V., Raynaud, L., Cohen, L., Mehler, J., & Mazoyer, B. (1996). Cerebral activations during number multiplication and comparison: A PET study. *Neuropsychologia, 34,* 1097-1106.

Del Abril, A., Segovia, S., & Guillamon, A. (1987). The bed nucleus of the stria terminalis in the rat: Regional sex differences controlled by gonadal steroids early after birth. *Brain Research, 429,* 295–300.

Del Abril, A., Segovia, S. & Guillamon, A. (1990). Sexual dimorphism in the parastrial nucleus of the rat preoptic area. *Developmental Brain Research, 52,* 11-15.

De Lacoste-Utamsing, C. D., & Holloway, R. L. (1982). Sexual dimorphism in the human corpus callosum. *Science, 216,* 1431-1432.

De Lacoste-Utamsing, C. D., Holloway, R. L., & Woodward, D. (1986). Sex differences in the fetal human corpus callosum. *Human Neurobiology, 5,* 93-96.

Delazer, M., & Benke, T. (1997). Arithmetic facts without meaning. *Cortex, 33,* 697-710.

Delgado, A. R., & Prieto, G. (2004) Cognitive mediator and sex-related differences in mathematics. *Intelligence, 32,* 25-32.

Deloche, G., von Aster, M. G., Dellatolas, G., Gaillard, F., Tiéche, C., & Azema, D. (1995). Traitement des nombres et calcul en CE 1 et CE 2. *Approche Neuropsychologique des Apprentissages chez l'Enfant (A.N.A.E.),* Hors Série, 42–52.

Department of Education and Skills (2002). (Data file). Available from the DfES statistics website retrieved June 4' 2006 from http://www.dfes.gov.uk/

Desoete, A., & Grégoire, J. (2007). Numerical competence in young children and in children with mathematics learning disabilities, *Learning and Individual Differences, 16*, 351–367.

De Vries, G. J. & Simerly, R. B. (2002). Anatomy, Development, and Function of Sexually Dimorphic Neural Circuits in the Mammalian Brain. In D. W. Pfaff, A. P. Arnold, A. M. Etgen, S. E. Fahrbach, R. L. Moss, & R. T. Rubin (Eds.) *Hormones,* *Brain, and Behavior. Volume IV. Development of Hormone-Dependent Neuronal Systems* (pp. 137-191). San Diego, U.S.A.: Academic Press.

Diamond, M. C., Johnson, R. E. & Ingham, C. A. (1975). Morphological changes in the young, adult and aging rat cerebral cortex, hippocampus and diencephalon. *Behavioral Biology, 14,*163–174.

Diamond, M. C., Johnson, R. E. & Ehlert, J. (1979). A comparison of cortical thickness in male and female rats—normal and gonadectomized, young and adult. *Behavioural and Neural Biology, 26,* 485–491.

Diamond, M. C., Dowling, G. A. & Johnson, R. E. (1981). Morphologic cerebral cortical asymmetry in male and female rats. *Experimental Neurology, 71,* 261–268.

Dohler, K. D., Hancke, J. L., Srivastava, S. S., Hofmann, C., Shryne, J. E. & Gorski, R. A. (1984). Participation of estrogens in female sexual differentiation of the brain; neuroanatomical, neuroendocrine and behavioral evidence. *Progress in Brain Research, 61,* 99–117.

Dubb, A., Gur, R., Avants, B., & Gee, J. (2003). Characterization of sexual dimorphism in the human corpus callosum. *NeuroImage, 20,* 512-519.

Ducharme, J. R., Forest, M. G., De Peretti, E., Sempe, M. & Bertrand, J. (1976). Plasma adrenal and gonadal sex steroids in human pubertal development. *Journal of Clinical Endocrinology & Metabolism, 42,* 468-476.

Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of reading and arithmetic skills in 7- to 10-year-olds. *Journal of Experimental Child Psychology, 91,* 113-136.

Eals, M., & Silverman, I. (1994). The hunter-gatherer theory of spatial sex differences: Proximate factors mediating the female advantage in recall of object arrays. *Ethology and Sociobiology, 15,* 95-105.

Eger, E., Sterzer, P., Russ, M. O., Giraud, A-L., Kleinschmidt, A. (2003). A supremodal number representation in human intraparietal cortex. *Neuron, 37,* 719-726.

Eglinton, E., & Annett, M. (1994). Handedness and dyslexia: A meta-analysis. *Perceptual and Motor Skills, 79,* 1611-1616.

Ehrhardt, A. A. & Baker, S. W. (1974). Fetal androgens, human central nervous system differentiation, and behaviour sex differences. In R. C. Richart, & R. L. Vande Wiele (Eds.) *Sex differences in behaviour* (pp. 33-51). New York: Wiley.

Einon, D. (1980). Spatial memory and response strategies in rats: Age, sex and rearing difference in performance. *Quarterly Journal of Experimental Psychology, 32,* 473-489.

Elkadi, S., Nicholls, M. E. R. & Clode, D. (1999). Handedness in opposite and same-sex dizygotic twins: Testing the testosterone hypothesis. *Neuroreport, 10,* 333-336.

Emmerton, J. (1998). Numerosity differences and effects of stimulus density on pigeons' discrimination performance. *Animal Learning & Behavior, 26,* 243-256.

Erlanger, D. M., Kutner, K. C. & Jacobs, A. R. (1999). Hormones and cognition: Current concepts and issues in neuropsychology. *Neuropsychological Review, 9,* 175-207.

Even, M. D., Dhar, M. G., & vom Saal, F. S. (1992). Transport of steroids between foetuses via amniotic fluid in relation to the intrauterine position phenomenon in rats. *Journal of Reproduction and Fertility, 96,* 709-716.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

Feigenson, L., & Carey, S. (2003) Tracking individuals via object-files: evidences from infants' manual search. *Developmental Science* **6**: 568-584

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology, 44*, 33-66.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307-314.

Fennema, E. (1974). Mathematics learning and the sexes. *Journal for Research in Mathematics Education, 5,* 126-129.

Field, A. (2009). *Discovering statistics using SPSS. Third Edition.* London: Sage.

Filipek, P. A., Richelme, C., Kennedy, D. N., & Caviness, V. S. (1994). The young adult human brain: An MRI-based morphometric analysis. *Cerebral Cortex, 4,* 344-360.

Finegan, J. K., Bartleman, B. & Wong, P. Y. (1989). A window for the study of prenatal sex hormone influences on postnatal development. *Journal of Genetic Psychology, 150,* 101-112.

Finegan, J, K., Niccols, G. A. & Sitarenios, G. (1992). Relations between prenatal testosterone levels and cognitive abilities at 4 years. *Developmental Psychology*, *28,* 1075-1089.

Fink, B., Manning, J.T., Neave, N., & Tan, U. (2004). Second to fourth digit ratio and hand skill in Austrian children. *Biological Psychology, 67,* 375-384.

Fink, B., Brookes, H., Neave, N., Manning, J. T., & Geary, D. C. (2006). Second to fourth digit ratio and numerical competence in children. *Brain and Cognition, 61,* 211-218.

Fitch, R. H, Cowell, P. E., Schrott, L. M. & Denenberg, V. H. (1991). Corpus callosum: ovarian hormones and feminization. *Brain Research, 542,* 313–317.

Flombaum, J. I., Junge, J. A., & Hauser, M. D. (2005). Rhesus monkeys (*Macaca mulatta*) spontaneously compute addition operations over large numbers. *Cognition, 97,* 315-325, Friedman, L. (1989). Mathematics and the gender gap: A meta-analysis of recent studies on sex differences in mathematical tasks. *Review of Educational Research, 59,* 185-213.

Friedman, L. (1995). The space factor in mathematics – Gender differences. *Review of Educational Research, 65,* 22-50.

Fuchs, A. R. & Fuchs, F. (1984). Endocrinology of human parturition: A review. *British Journal of Obstetrics and Gynaecology, 91,* 948-967.

Gallagher, A. & DeLisi, R. (1994). Gender differences in Scholastic Aptitude Test-Mathematics problem solving among high-ability students. *Journal of Educational Psychology, 86,* 204-211.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43-74.

Garcia-Falgueras, A., Pinos, H., Collado, P., Pasaro, E., Fernandez, R., Jordan C. J., Segovia, S. & Guillamon, A. (2005). The role of the androgen receptor in CNS masculinization. *Brain Research, 1035,* 13-23.

Garn, S. M., Burdi, A. R., & Babler, W. J. (1975). Early prenatal attainment of adult metacarpal-phalangeal rankings and proportions. *American Journal of Physical Anthropology, 43,* 327-332.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37,* 4-15. *Psychological Bulletin, 114,* 345-362.

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist, 50,* 24–37.

Geary, D. C. (1996). Sexual selection and sex differences in mathematical ability. *Behavioural and Brain Science, 19*, 229-284.

Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology, 77*, 236–263.

George, R. (1930). Human finger types. *Anatomical Record, 46*, 199-204.

Geschwind, N., & Galaburda, A. M. (1987). *Cerebral Lateralization: Biological mechanisms, associations and pathology.* Cambridge, MA: MIT Press.

Gevers, W., Reynvoet, B., & Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition, 87*, B87–B95.

Gevers, W., Reynvoet, B., & Fias, W. (2004). The mental representation of ordinal sequences is spatially organized. Evidence from days of the week. *Cortex, 40*, 171–172.

Gilbert, A. N., & Wysocki, C. J. (1992). Hand preferences and age in the United States. *Neuropsychologia, 30*, 601-608.

Gobrogge, K., Breedlove, S. M., & Klump, K. L. (2008). Genetic and environmental influences on 2D:4D finger length ratios: A study of monozygotic and dizygotic male and female twins. *Archives of Sexual Behavior, 37*, 112-118.

Gooren, L. (2007). Testosterone and the brain. *The journal of men's health & gender, 4,* 344-351.

Gordon, H. W., & Lee, P. A. (1986). A relationship between gonadotropins and visuospatial function. *Neuropsychologia, 24*, 563-576.

Gorski, R. A. (1984). Critical role for the medial preoptic area in the sexual differentiation of the brain. In G. J. De Vries, J. P. C. De Bruin, H. B. M. Uylings, A. Corner (Eds.) *Sex differences in the brain, Progress in brain research: Volume 61* (pp.129-146) Amsterdam: Elsevier.

Gorski, R. A. (2002). Hypothalamic imprinting by gonadal steroid hormones. *Advances in Experimental Medicine and Biology, 511*, 51-70.

Gorski, R. A., Gordon, J. H., Shryne, J. E. & Southam, A .M. (1978). Evidence for a morphological sex difference within the medial preoptic area of the rat brain. *Brain Research, 148,* 333-346.

Gorski, R. A., Harlan, R. A., Jacobson, C. D., Shryne, J. E. & Southam, A. M. (1980). Evidence for existence of a sexually dimorphic sub-nucleus in the preoptic area of the rat. *The Journal of Comparative Neurology, 193,* 529-539.

Gouchie, C., & Kimura, D. (1991). The relationship between testosterone levels and cognitive ability patterns. *Psychoneuroendocrinology, 16,* 323-334.

Goy, R. W., & McEwen, B. S. (1980). *Sexual differentiation of the brain.* Massachusetts, U.S.A.: MIT Press.

Grimshaw, G. M., Bryden, M. P. & Finegan, J. K. (1995). Relations between prenatal testosterone and cerebral lateralization in children. *Neuropsychology, 9,* 68-79.

Grimshaw, G. M., Sitarenios, G., & Finegan, J. K. (1995). Mental rotation at 7 years: Relations with prenatal testosterone levels and spatial play experiences. *Brain and Cognition, 29*, 85-100.

Grogger, J., & Eide, E. (1995). Changes in college skills and the rise in the college wage premium. *Journal of Human Resources, 30,* 280-310.

Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). Developmental dyscalculia: Prevalence and demographic features. *Developmental Medical Child Neurology, 38,* 25-33.

Guillamon, A., De Blas, M. R. & Segovia, S. (1988). Effects of sex steroids on the development of the locus coeruleus in the rat. *Developmental Brain Research, 40,* 306–310.

Gur, R. C., & Gur, R. E. (2004). Gender differences in the functional organization of the brain. In M. J. Legato (Ed.) *Principles of gender-specific medicine* (pp. 63-70). Amsterdam: Elsevier.

Gur, R. C., Turetsky, B. I., Matsui, M., Yan, M., Bilker, W., Hughett, P., & Gur, R. E. (1999). Sex differences in brain gray and white matter in healthy young adults: Correlations with cognitive performance. *Journal of Neuroscience, 19, I4065-4072.*

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (2005). The neuroanatomy of general intelligence: sex matters. *NeuroImage, 25,* 320-327.

Halpern, D. F. (2000). *Sex differences in cognitive abilities. Third edition.* New Jersey, U.S.A.: Lawrence Erlbaum Associates Inc.

Halpern, C., Clark, R., Moore, P., Cross, K., & Grossman, M. (2007). Too much to count on: Impaired very small numbers in corticobasal degeneration. *Brain and Cognition, 64,* 144-149.

Hamilton, C. (2008). *Cognition and Sex Differences.* Hampshire: Palgrave Macmillan.

Hampson, E. (1995). Spatial cognition in humans: Possible modulation by androgens and estrogens. *Journal of Psychiatry and Neuroscience, 20,* 397-404.

Hampson, E., Ellis, C. L., & Tenk, C. M. (2008). On the relation between 2D:4D and sex-dimorphic personality traits. *Archives of Sexual Behaviour, 37,* 133-144.

Hampson, E., & Kimura, D. (1988). Reciprocal effects of hormonal fluctuations on human motor and perceptual-spatial skills. *Behavioral Neuroscience, 102,* 456-459.

Hampson, E., Rovet, J. F. & Altmann, D. (1998). Spatial reasoning in children with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Developmental Neuropsychology, 14,* 299-320.

Hardie, D. G. (1991). *Biochemical messengers: Hormones, Neurotransmitters and growth factors*. Cambridge, U.K.: Chapman & Hall.

Hassler, M., & Gupta, D. (1993). Functional brain organisation, handedness and immune vulnerability in musicians and non-musicians. *Neuropsychologia, 31,* 655–660.

Hauser, M. D. (2000). *Wild minds: What animals really think.* New York: Henry Holt, Inc.

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society B: Biological Sciences, 267,*829–833.

Hauser, M. D., & Carey, S. (2003). Spontaneous representations of small numbers of objects by rhesus macaques: Examinations of content and format. *Cognitive Psychology*, *47,* 367-401.

Hauser, M. D. & Spelke, E. (2004). Evolutionary and developmental foundations of human knowledge: A case study of mathematics. In M. D. Gazzaniga *The cognitive neurosciences* (pp 853-864). Cambridge, Massachusetts: MIT Press.

Hauser, M. D., Tsao, F., Garcia, P., & Spelke, E. S. (2003). Evolutionary foundations of number: spontaneous representation of numerical magnitudes by cotton-top tamarins. *Proceedings of the Royal Society of London, 270,* 1441-1446.

Hausmann, M., Slabbekoorn, D., Van Goozen, S.H., Cohen-Kettenis, P.T. & Gunturkun, O. (2000). Sex hormones affect spatial abilities during the menstrual cycle. *Behavioral Neuroscience, 114,* 1245–1250.

Helleday, J., Bartfai, A., Ritzen, E. M., & Forsman, M. (1994). General intelligence and cognitive performance in women with congenital adrenal hyperplasia (CAH). *Psychoneuroendocrinology, 19,* 343-356.

Hellige, J. B., & Longstreth, L. E. (1981) Effects of concurrent hemisphere specific activity on unimanual tapping rate. *Neuropsychologia, 19,* 395-405.

Hendricks, S. E. (1992). Role of estrogen and progestins in the development of female sexual behaviour potential. In A. A. Gerall, H. Moltz, & I. L. Ward (Eds.) *Sexual differentiation: Handbook of behavioural neurobiology* (pp. 129-155). New York, U.S.A.: Pleum.

Henschen, S. E. (1919). Über Sprach- Musil- und Rechenmechanismen und ihre Lokalisationen im Grosshirn. *Zeitschrift für die desamte Neurologie und Psychiatrie, 52*, 273-298.

Herman, R. A. & Wallen, K. (2007). Cognitive performance in rhesus monkeys varies by sex and prenatal androgen exposure. *Hormones and Behaviour, 51,* 469-507.

Hier, D. B. & Crowley, W. F., Jr. (1982). Spatial ability in androgen deficient men. *New England Journal of Medicine, 306*, 1202-1205.

Hines, M., Ahmed, S. F. & Hughes, I. A. (2003). Psychological outcomes and gender-related development in complete androgen insensitivity syndrome. *Archives of Sexual Behaviour*, *32,* 93-101.

Hines, M., Fane, B. A., Pasterski, V. L., Mathews, G. A., Conway, G. S., & Brook, C. (2003). Spatial abilities following prenatal androgen abnormality: Targeting and mental rotations performance in individuals with congenital adrenal hyperplasia. *Psychoneuroendocrinology, 28,* 1010-1026.

Hines, M., & Shipley, C. (1984) Prenatal exposure to diethylstilbestrol (DES) and the development of sexually dimorphic cognitive abilities and cerebral lateralisation. *Developmental Psychology, 20,* 81-94.

Holt, S. B. (1968). *The genetics of dermal ridges*. Springfield, IL: Charles C. Thomas.

Hönekopp, J., Bartholdt, L., Beier, L., & Liebert, A. (2007). Second to fourth digit length ratio (2D:4D) and adult sex hormone levels: New data and a meta-analytic review. *Psychoneuroendocrinology, 32,* 313-321.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: a meta-analysis. *Psychological Bulletin, 107,* 139-155

Imperato-McGinley, J., Pichardo, M., Gautier, T., Voyer, D. & Bryden, M. P. (1991). Cognitive abilities in androgen-insensitive subjects: Comparison with control males and females from the same kindred. *Clinical Endocrinology, 34,* 341-347.

Isaacs, K. L., Barr, W. B., Nelson, P. K., & Devinsky, O. (2006). Degree of handedness and cerebral dominance. *Neurology, 66*, 1855-1858.

Isaacs, E. B., Edmonds, C. J., Lucas, A., & Gadian, D. G. (2001). Calculation difficulties in children of very low birth weight: A neural correlate. *Brain, 124,* 1701-1707.

Isgor, C. & Sengelaub, D. R. (1998). Prenatal gonadal steroids affect adult spatial behavior, CA1 and CA3 pyramidal cell morphology in rats. *Hormones and Behaviour, 34,* 183-198.

Jacklin, C. N., Wilcox, K. T., & Maccoby, E. E. (1988). Neonatal sex-steroid hormones and cognitive abilities at six years. *Developmental Psychobiology, 21,* 567-574.

Jackson, N., & Coney, J. (2004). Right hemisphere superiority for subitizing. *Laterality, 9,* 53-66.

Jacobson, C. D., Csernus, V. J. & Gorski, R. A. (1981). The influence of gonadectomy, androgen exposure, or a gonadal graft in the neonatal rat on the volume of the sexually dimorphic nucleus of the preoptic area. *Journal of Neuroscience, 1,* 1142-1147.

Jacobsen, C. D., Csernusm V. J., Shryne, J. E., & Gorski, R. A. (1981). The influence of gonadectomy, androgen exposure, or a gonadal graft in the neonatal rat on the volume of the sexually dimorphic nucleus of the preoptic area. *Journal of Neuroscience, 1,* 1142-1147.

Jensen, A. R. (1988). Sex differences in arithmetic computation and reasoning in prepubertal boys and girls. *Behavioral and Brain Sciences, 11,* 198-199.

Jones, B. A. & Watson, N. V. (2005). Spatial memory performance in androgen insensitive male rats. *Physiology & Behviour, 85,* 135-141.

Jordan, K. E., & Brannon, E. M. (2006). The multisensory representation of number in infancy. *Proceedings of the National Academy of Sciences, 103,* 3486-3489.

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with species mathematics difficulties versus children with co morbid mathematics and reading difficulties. *Child Development, 74,* 834–850

Jordan, N. C., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77,* 153-175.

Joseph, R., Hess, S. & Birecree, E. (1978). Effects of hormone manipulations and exploration on sex differences in maze learning. *Behavioural Biology, 24,* 364-377.

Judd, H. L., Robinson, J. D., Young, P. E. & Jones, O. W. (1976). Amniotic fluid testosterone levels in midpregnancy. *Obstetrics & Gynecology, 48,* 690-692.

Kallai, J., Csathó, A., Kövėr, F., Makány, T., Nemes, J., Horváth, K., Kovács, N., Manning, J. T., Nadel, L., & Nagy, F. (2005). MRI-assessed volume of left and right hippocampi in females correlates with the relative length of the second and fourth fingers (the 2D:4D ratio). *Psychiatry Research: Neuroimaging, 140,* 199-210.

Kanit, L., Taskiran, D., Furedy, J. J., Kulali, B., McDonald, R. & Pogun, S. (1998). Nicotine interacts with sex in affecting rat choice between “look-out” and “navigational” cognitive styles in the Morris water maze place learning task. *Brain Research Bulletin, 46,* 441–445.

Kaufman, L. (2002). More evidence for the role of the central executive in retrieving arithmetic facts: A case study of severe developmental dyscalculia. *Journal of Clinical and Experimental Neuropsychology, 24,* 302-310.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Wolkmann, J. (1949). The discrimination of visual number. *The American Journal of Psychology, 62,* 498-525.

Kawata, M. (1995). Roles of steroid hormones and their receptors in structural organization in the nervous system. *Neuroscience Research, 24,* 1-46.

Keenan, P. A. & Soleymani, R. M. (2001). Gonadal steroids and cognition. In R. Tarter, & M. Butters (Eds.) *Medical neuropsychology: The impact of disease on behaviour. Critical Issues in Neuropsychology* (pp. 181-197). New York, U.S.A.: Plenum Press.

Kempel, P., Gohlke, B., Klempau, J., Zinsberger, P., Reuter, M., & Hennig, J. (2005). Second to fourth digit length, testosterone and spatial ability. *Intelligence, 3,* 215-230.

Kertzman, C., Robinson, D. L. Sherins, R. J., Schwankhaus, J. D. & McClurkin, J. W. (1990). Abnormalities in visual spatial attention in men with mirror movements associated with isolated hypogonadotropic hypogonadism. *Neurology, 40*, 1057-1063.

Kiefer, M., & Dehaene, S. (1997). The time course of parietal activation in single-digit multiplication: Evidence from event related potentials. *Mathematical Cognition, 3,* 1-30.

Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin, 105,* 198-214.

Kimura, D. (1994). Body asymmetry and intellectual pattern. *Personality and Individual Differences, 19,* 471-478.

Kimura, D. (1996). Sex, sexual orientation and sex hormones influence human cognitive function. *Current Opinions in Neurobiology, 6,* 259-63.

Kimura, D. (1999). *Sex and Cognition*. Masseuses, U.S.A.: MIT Press.

Kimura, D., & Carson, M. W. (1995). Dermatoglyphic asymmetry: relation to sex, handedness, and cognitive pattern. *Personality and Individual Differences, 19,* 471–478.

Kimura, D., & Clarke, P. G. (2001). Cognitive pattern and dermatoglyphic asymmetry. *Personality and Individual Differences, 30,* 579-586.

Knecht, S., Drager, B., Deppe, M., Bobe, L., Lohmann, H., Floel, A., et al. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain, 123,* 2512-2518.

Knickmeyer, R. C. & Baron-Cohen, S. (2006). Fetal testosterone and sex differences. *Early Human Development, 82,* 755-760.

Koechlin, E. (1997). Numerical Transformations in Five-month-old Human Infants. *Mathematical Cognition, 3,* 89-104.

Kolata, G. (1983). Math genius may have a hormonal basis. *Science, 222,* 1312-1326.

Kondo, T., Zakany, J., Innis, W. J., & Duboule, D. (1997). Of fingers, toes and penises. *Nature, 390,* 29.

Koontz, K. L., & Berch, D. B. (1996). Identifying simple numerical stimuli: Processing inefficiencies exhibited by arithemetic learning disabled children. *Mathematical Cognition, 2*, 1-23.

Kosc, L. (1974). Developmental dyscalculia. *Journal of Learning Disabilities, 7,* 164-177.

Kucian, K., Loenneker, T., Dietrich, T., Dosch, M., Martin, E., & von Aster, M. (2006). Impaired neural networsd for approximate calculation in dyscalculic children. *Behavioral and Brain Functions, 2,* 31.

Kyttälä, M., Aunio, P., Lehto, J. E., Van Luit, J., & Hautamäki, J. (2003). Visuospatial working memory and early numeracy. *Educational and Child Psychology, 20,* 65-76.

Landerl, K., Bevan A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8–9-year-old students. *Cognition, 93,* 99-125.

Langdon, D. W., & Warrington, E. K. (1997). The abstraction of numerical relations: a role for the right hemisphere in arithmetic? *Journal of the International Neuropsychological Society, 3,* 260-268.

Lansky, L., Feinstein, H., & Peterson, J. M. (1988). Demography of handedness in two samples of rendomly selected adults (N=2083). *Neuropsychologia, 26,* 465-477.

Leahey, E., & Guo, G. (2001). Gender differences in mathematical trajectories. *Social Forces, 80,* 713-732.

Leret, M. L., Molina-Holgado, F. & Gonzalez, M. I. (1994). The effect of perinatal exposure to estrogens on the sexually dimorphic response to novelty. *Physiology and Behaviour, 55,* 371–373.

Levy, J. (1969). Possible basis for the evolution of lateral specialization in the human brain, *Nature, 224,* 614–615.

Levy, J., & Reid, M. (1978). Variations in cerebral organisation as a function of handedness, hand posture in writing and sex. *Journal of Experimental Psychology: General, 107,* 119-144.

Levy, J., & Reis, I. L. & Grafman, J., (1999). Metabolic abnormalities detected by H-MRS in dyscalculia dysgraphia. *Neurology, 53,* 639-641.

Lewis, C., Hitch, G. J., & Walker, P. (1994). The prevalence of specific arithmetic difficulties and specific reading difficulties in 9 to 10 year old boys and girls. *Journal of Child Psychology and Psychiatry, 35,* 283-292.

Linn, M. C. & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta analysis. *Child Development, 56,* 1479-1489.

Lipton, S. S., & Spelke, E. S. (2003). Origins of number sense: Large number discrimination in human infants. *Psychological Science, 15,* 396-401.

Loehlin, J. C. & McFadden, D. (2003). Otoacoustic emissions, auditory evoked potentials, and traits related to sex and sexual orientation. *Archives of Sexual Behaviour, 32,* 115-127.

Luder, E., Steinmetz, H., & Jancke, L. (2002). Brain size and grey matter volume in the healthy human brain. *Neuroreport, 13,* 2371-2374.

Lutchmaya, S., Baron-Cohen, S. & Raggatt, P. (2002) Foetal testosterone and vocabulary size in 18- and 24-month-old infants. *Infant Behaviour and Development, 24,* 418-424.

Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., & Manning, J. T. (2004). 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Human Development*, *77,* 23-28.

Luxen, M. F., & Buunk, B. P. (2005). Second to fourth digit ratio related to verbal and numerical intelligence and the big five. *Personality and Individual Differences*, *39,* 959-966.

Maccoby, E. E., Doering, C. H., Jacklin, C. N. & Kraemer, H. (1979). Concentrations of sex hormones in umbilical cord blood: Their relation to sex and birth order of infants. *Child Development, 50,* 632-642.

Maccoby, E. E. & Jacklin, C. N. (1974). *The Psychology of Sex Differences*. Stanford, U.S.A.: Stanford University Press.

Machado, A., & Keen, R. (2002). Relative numerostiy discrimination in the pigeon: further tests of the linear exponential ration model. *Behavioural Processes, 57,* 131-148.

Mandler, G., & Shebo, B. J. (1982) Subitizing: an analysis of its component processes. *Journal of Experimental Psychology: General, 111,* 1-21

Manning, J. T. (2002). *Digit ratio: a pointer to fertility, behaviour and health.* New Brunswick: Rutgers University Press.

Manning, J. T., Barley, L., Lewis-Jones, I., Walton, J., Trivers, R. L., Singh, D., Thornhill, R., Rohde, P., Bereczkei, T., Henzi, P., Soler, M., & Szwed, A. (2000). The 2nd to 4th digit ratio, sexual dimorphism, population differences and reproductive success: Evidence for sexual antagonistic genes? *Evolution and Human Behaviour, 21,*163-183.

Manning, J. T., & Bundred, P. E. (2000). The ratio of 2nd to 4th digit length: a new prediction of disease predisposition? *Medical Hypothesis, 54,* 855-857.

Manning, J. T., Bundred, P. E., Newton, D. J., & Flanagan, B. F. (2003). The second to fourth digit ratio and variation in the androgen receptor gene. *Evolution and Human Behaviour, 24*, 399-405.

Manning, J. T., Fink, B., Neave, N., & Caswell, N. (2005). Photocopies yield lower digit rations (2D:4D) than direct finger measurements. *Archives of Sexual Behavior, 34*, 239-333.

Manning, J. T., Henzi, P., Venkatramana, P., Martin, S. & Singh, D. (2003). Second to fourth digit ratio: Ethnic differences and family size in England, Indian and South African populations. *Annals of Human Biology, 30,* 579-588.

Manning, J. T., & Peters, M. (2009). Digit ration (2D:4D) and hand preference for writing in the BBC internet study. *Laterality: Asymmetries of Body, Brain and Cognition, 14,* 528-540.

Manning, J. T., Scutt, D., Wilson, J., & Lewis-Jones, D. I. (1998). The ratio of 2nd to 4th digit length: a predictor of sperm numbers and concentrations of testosterone, luteinizing hormones and oestrogen. *Human Reproduction, 13,* 300-3004.

Manning, J. T., Stewart, A., Bundred, P. E. & Trivers, R. L. (2004). Sex and ethnic differences in 2nd to 4th digit ratio of children. *Early Human Development, 80,* 161-168.

Manning, J. T. & Taylor, R. P. (2001). 2nd to 4th digit ratio and male ability in sport: Implications for sexual selection in humans. *Evolution and Human Behaviour, 22,* 61-69.

Manning, J. T., Trivers, R. L., Singh, D., & Thornhill, R. (1999). The mystery of female beauty. *Nature, 399,* 214-215.

Manning, J. T., Trivers, R. L., Thornhill, R. & Singh, D. (2000). The 2^{nd} :4^{th} digit ratio and asymmetry of hand performance in Jamaican children. *Laterality, 5,* 121-132.

Manning, J. T., Wood, S., Vang, E., Walton, J., Bundred, P. E., van Heyningen, C. & Lewis-Jones, D. I. (2004). Second to fourth digit ratio (2D:4D) and testosterone in men. *Asian Journal of Andrology, 6,* 211–215.

Matsumoto, A. & Arai, Y. (1983). Sex differences in volume of the ventromeidal nucleus of the hypothalamus in the rat. *Endocrinologia japonica, 30,* 277-280.

Matsumoto, A. & Arai, Y. (1986). Male-female differences in synaptic organization of the ventromedial nucleus of the hypothalamus in the rat. *Neuroendocrinology, 42,* 232-236.

Mayringer, H., & Wimmer, H. (2002). No deficits at point of hemispheric indecision. *Neuropsychologia, 40,* 701–704.

McBurney, D. H., Gaulin, S. J. C., Devineni, T., & Adams, C. (1997). Superior spatial memory of women: Stronger evidence for the gathering hypothesis. *Evolution and Human Behavior, 18,* 165-174.

McCrink, M., & Wynn, K. (2004). Large number addition and subtraction by 9-month-old infants. *Psychological Science, 15,* 776-781.

McFadden, D. (1998). Sex differences in the auditory system. *Developmental Neuropsychology, 14,* 261-298.

McFadden, D. (1993). A masculinising effect on the auditory systems of human females having male co-twins. *Proceedings of the National Academy of Sciences USA, 90,* 11900-11904.

McFadden, D. (2002). Masculinization effects in the auditory system. *Archives of Sexual Behaviour, 31,* 99-111.

McFadden, D. (2000). Masculinising effects on otoacoustic emissions and auditory evoked potentials in women using oral contraceptives. *Hearing Research, 142,* 23–33.

McFadden, D. & Champlin, C. A. (2000). Comparison of auditory evoked potentials in, heterosexual, homosexual, and bisexual males and females. *Journal of the Association for research in Otolaryngology, 1,* 89-99.

McFadden, D. & Loehlin, J. C. (1995). On the heritability of spontaneous otoacoustic emissions: A twins study. *Hearing Research, 85,* 181-198.

McFadden, D., Loehlin, J. C. & Pasanen, E. G. (1996). Additional findings on heritability and prenatal masculinisation of cochlear mechanisms: Click-evoked otoacoustic emissions. *Hearing Research, 97,* 102-119.

McFadden, D. & Pasanen, E. G. (1998). Comparison of the auditory systems of heterosexuals and homosexuals: Click-evoked otoacoustic emissions. *Proceedings of the National Academy of Sciences USA, 95,* 2709-2713.

McFadden, D. & Pasanen, E. G. (1999). Spontaneous otoacoustic emissions in heterosexuals, homosexuals, and bisexuals. *Journal of the Acoustic Society of America, 105,* 2403-2413.

McFadden, D. & Schubel, E. (2002). Relative lengths of fingers and toes in human males and females. *Hormones and Behaviour, 42,* 492-500.

McFadden, D., & Schubel, E. (2003). The relationship between otoacoutic emissions and relative lengths of fingers and toes. *Hormones and Behaviour, 43,* 421–429.

McGowan, J. G., & Duka, T. (2000) Hemispheric lateralisation in a manual task-verbal task combination: the role of modality and gender. *Neuropsychologia, 38,* 1018-1027

McGuire, L. S., Ryan, K. O. & Omenn, G. S. (1975). Congenital adrenal hyperplasia. II. Cognitive and behavioural studies. *Behaviour Genetics, 5,* 175-188.

McIntosh, S., & Vignoles, A. (2001). Measuring and assessing the impact of basic skills on labour market outcomes. *Oxford Economic Papers, 53,* 453-481.

McIntyre, M. H. (2006). The use of digit ratios as markers for perinatal androgen action. *Reproductive Biology and Endocrinology, 4,* 10-18.

McIntyre, M. H., Cohn, B. A., & Ellison, P. T. (2006). Sex dimorphism in digital formulae of children. *American Journal of Physical Anthropology, 129,* 143-50.

McIntyre, M. H., Ellison, P. T., Lieberman, D. E., Demerath, E. & Towne B. (2005). The development of sex differences in digital formula from infancy in the Fels Longitudinal Study. *Proceedings of the Royal Society: B, 272,* 1473–1479.

McKeever, W. F., & Deyo, R. A. (1990). Testosterone, dihydrotestosterone, and spatial task performance of males. *Bulletin of the Psychonomic Society, 28,* 305-308.

McKeever, W. F., Rich, D. A., Deyo, R. A., & Conner, R. L. (1987). Androgens and spatial ability: Failure to find a relationship between testosterone and ability measures. *Bulletin of the Psychonomic Society, 25,* 438-440.

McManus, I. C., & Mascie-Taylor, C. G. N. (1983). Biosocial correlates of cognitive abilities. *Journal of Biosocial Science, 15,* 289–306.

McManus, I. C., Murray, B., Doyle, K., & Baron-Cohen, S. (1992). Handedness in childhood autism shows a dissociation of skill and preference, *Cortex, 28,* 373-381.

Meck W.H., & Church, R.M. (1983). A mode control model of counting and timing processes. *Journal Experimental Psychology: Animal Behaviour Process, 9,* 320–334.

Medland, S. E., & Loehlin, J. C. (2008). Multivariate Genetic Analyses of the 2D:4D Ratio: Examining the Effects of Hand and Measurement Technique in Data from 757 Twin Families. *Twin Research and Human Genetics, 11,* 335-341.

Mennon, V., Rivera, S. M., White, C. D., Glover, G. H., & Reiss, A. L. (2000). Dissociating prefrontal and parietal cortex activation during arithmetic processing. *Neuroimage, 12,* 357-365.

Miller, E. M. (1994). Prenatal sex hormone transfer: A reason to study opposite-sex twins. *Personality and Individual Differences, 17,* 111-529.

Miller, G. M., & Chapman, J. P. (1987). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110,* 40-48.

Miller, W. L. & Levine, L. S. (1987). Molecular and clinical advances in congenital adrenal hyperplasia. *The Journal of Pediatrics, 111,* 1–17.

Mittwoch, U. & Mahadevaiah, S. (1980). Additional growth: A link between mammalian testes, avian ovaries, gonadal asymmetry in hermaphrodites and the expression of H-Y antigen. *Growth, 44,* 287-300.

Mizukami, S., Nishizuka, M. & Arai, Y. (1983). Sexual difference in nuclear volume and its ontogeny in the rat amygdala. *Experimental neurology, 79,* 569-575.

Moffat, S. D., & Hampson, E. (1996). A curvilinear relationship between testosterone and spatial cognition in humans: Possible influence of hand preference. *Psychoneuroendocrinology, 21,* 232-337.

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature, 215,* 1519-1520.

Muller, C. (1998). Gender differences in parental involvement and adolescents' mathematics achievement. *Sociology of Education, 71,* 336-356.

Műller, M. J. (1994). Salivary testosterone and simple reaction time parameters. *Neuropsychobiology, 30,* 173-177.

Mullis, I. V. S., Martin, M., O., Gonzales, E. J., Gregory, E. J., Garden, R. A., O'Connor, A. M., et al. (1999). *Findings from IEA's repeat of the Third International Mathematics and Science Study and the eighth grade.* Boston: International Study Centre.

Nan, Y., & Knösche, T. R., & Luo, Y-J. (2005). Counting in everyday life: Discrimination and enumeration. *Neuropsychologia, 44,* 1103-1113.

Nass, R., Baker, S., Speiser, P., Virdis, R., Balsamo, A., Caccian, E., Loche, A., Dumic, M. & New, M. (1987). Hormones and handedness: Left-hand bias in female congenital adrenal hyperplasia patients. *Neurology, 37,* 711-715.

National Centre for Education and Statistics (2004). Digest of education statistics 2004 (Data file). Available from National Centre for Education Statistics web site, retrieved June 20, 2006 from http://www.nces.ed.gov

Neave, N. (2008). *Hormones and behaviour: A psychological approach.* Cambridge: Cambridge University Press.

Neave, N., Laing, S., Fink, B. & Manning, J. T. (2003). Second to fourth digit ratio, testosterone and perceived male dominance. *Proceedings of the Royal Society of London Series B: Biological Sciences, 270,* 2167–2172.

Neave, N., Menaged, M., & Weightman, D. R. (1999). Sex differences in cognition: The role of testosterone and sexual orientation. *Brain and Cognition, 41,* 245-262.

Nelson, R. J. (2000). *An introduction to behavioural endocrinology: Second Edition*. Massachusetts, U.S.A.: Sinauer Associates, Inc.

Nettle, D. (2003). Hand laterality and cognitive ability: A multiple regression approach. *Brain and Cognition, 52,* 390-398.

Newcombe, F. G., Ratcliff, G. G., Carrivick, P. J., Hiorns, R. W., Harrison G. A., & Gibson J. B. (1975). Hand preference and IQ in a group of Oxfordshire villages. *Annals of Human Biology, 2,* 235–242.

Nicholls, M. E. R., Orr, C. A., Yates, M. J. & Loftus, A. M. (2008). A new means of measuring index/ring finger (2D:4D) ratio and its association with gender and hand preferences. *Laterality: Asymmetries or Body, Brain and Cognition, 13,* 71-91.

Nieder, A., & Miller, E. (2003). Coding of cognitive magnitude compressed scaling of numerical information in the primate prefrontal cortex. *Neuron, 37,* 149 – 157.

Nieder, A., & Miller, E. (2004). A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences, 101,* 7457-7462.

Nishizuka, M. & Arai, Y. (1981). Sexual dimorphism in synaptic organisation in the amygdala its dependence on neonatal hormone environment. *Brain Research, 212,* 31-38.

Nishizuka, M. & Arai, Y. (1981). Organizational action of estrogen on synaptic pattern in the amygdala: implications for sexual differentiation of the brain. *Brain Research, 214,* 422-426.

Nopoulos, P., Flaum, M., O'Leary, D., & Andreasen, N. C. (2000). Sexual dimorphism in the human brain: Evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. *Psychiatry Research, 98,* 1-13.

Norman, A. W., & Litwack, G. (1997). *Hormones: Second edition.* California, U.S.A.: Academic Press.

Noroozian, M., Lotfi, J., & Gassemzadeh, H., Emami, H., & Mehrabi, Y. (2002). Academic achievement and learning abilities in left-handers: Guilt or gift? *Cortex, 38,* 779–785.

O'Boyle, M. W., Alexander, J. E., & Benbow, C. P. (1991). Enhanced right hemisphere activation in the mathematically precocious: A preliminary EEG investigation. *Brain and Cognition, 17,* 138-153.

O'Boyle, M.W, & Benbow, C.P. (1990). Enhanced right hemisphere involvement during cognitive processing may relate to intellectual precocity. *Neuropsychologia, 28,* 211–216.

O'Boyle, M. W., Gill, H. S., Benbow, C. P., & Alexander, J. E. (1994). Concurrent finger-tapping in mathematically gifted males: Evidence for enhanced right hemisphere involvement during linguistic processing. *Cortex*, *30,* 519-526.

Ohno, S., Nakajima, Y., & Nakajin, S. (2005). Triphenytin and tributyltin inhibit pig testicular 17β-hydroxysteroid dehydrogenase activity and suppress testicular testosterone biosynthesis. *Steroids, 70,* 645-651.

Ökten, A., Kalyoncu, M., & Yaris, N. (2002). The ratio of second- and fourth-digit lengths and congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Early Human Development*, *70,* 47–54.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia, 9,* 97-113.

Palacios, S. (2007). Androgens and female sexual function. *Maturitas, 57,* 61-65.

Pang, S. (1997). Congenital adrenal hyperplasia. *Endocrinology and metabolism clinics of North America, 26,* 853-891.

Pang, S. & Shook, M. K. (1997). Current status of neonatal screening for congenital adrenal hyperplasia. *Current Opinion in Pediatrics, 9,* 419–423.

Pasini, M., & Tessari, A. (2001). Hemispheric specialization in quantification processes. *Psychological Research, 65,* 57-63.

Paterski, V., Hindmarsh, P., Geffner, M., Brook, C., Brain, C. & Hines, M. (2007). Increased aggression and activity level in 3- to 11-year-old girls with congenital adrenal hyperplasia (CAH). *Hormones and Behaviour,* 52, 368-374.

Paul, S. N., Kato, B. S., Cherkas, L. F., Andrew, T., & Spector, T. D. (2006). Heritability of the second to fourth digit ratio (2D:4D): A twin study. Twin Research and Human Genetics, 9, 215–219.

Peters, M., Reimers, S., & Manning, J. T. (2006). Hand preference for writing and associations with selected demographic and behavioural variables in 255,100 subjects: The BBC internet study. *Brain and Cognition, 62,* 177-189.

Perlman, S. M. (1973). Cognitive abilities of children with hormone abnormalities: Screening by psychoeducational test. *Journal of Learning Disabilities, 6,* 21-29.

Pesenti, M., Thioux, M., Seron, X., & De Volder, A. (2000). Neuroanatomical substrates of Arabic number processing, numerical comparison, and simple addition: A PET study. *Journal of Cognitive Neuroscience, 12*, 461-479.

Peters, M., Manning, J. T. & Reimers, S. (2007). The effects of sex, sexual orientation, and digit ratio (2D:4D) on mental rotation performance. *Archives of Sexual Behaviour, 36,* 251-260.

Peters, M., Tan, U., Kang, Y., Teixeira, L. & Mandal, M. (2002). Sex-specific finger length patterns linked to behavioural variables: Consistency across various human populations. *Perceptual and Motor Skills, 94,* 171181.

Phelps, V. R. (1952). Relative index finger length as a sex influenced trait in a man. *American Journal of Human Genetics, 4,* 72-89.

Phoenix, C. H., Goy, R. W., Gerall, A. A. & Young, W. C. (1959). Organising action of prenatally administered testosterone propionate on the tissues mediating mating behaviour in the female guinea pig. *Endocrinology, 65,* 369-382.

Piazza, M., Giacomini, E., Le Bihan, D., & Dehaene, S., (2003). Single-trial classification of parallel pre-attentive and serial attentive processes using functional magnetic resonance imaging. *Proceedings of the Royal Society of London Series B: Biological Sciences, 270,* 1237-1245.

Piazza, M., Mechelli, A., Butterworth, B., & Price, C. J. (2002). Are subitizing and counting implemented as separate or functionally overlapping processes? *Neuroimage, 15,* 435-446.

Pinel, P., Dehaene, S., Riviere, D., & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *NeuroImage, 14,* 281-288.

Pinos, H., Collado, P., Rodriguez-Zafra, M., Rodriguez, C., Segovia, S. & Guillamon, A. (2001). The development of sex differences in the locus coeruleus of the rat. *Brain Research Bulletin, 56,* 73–78.

Poulin, M., O'Connell, R., & Freeman, L. M. (2004). Picture recall skills correlate with 2D:4D ratio in women but not men. *Evolution and Human Behavior, 25,* 174-181.

Probst, R., Lonsbury-Martin, B. L. & Martin, G. K. (1991). A review of otoacoustic emissions. *The Journal of the Acoustical Society of America, 89,* 2027-2067.

Prinzel, L. J. & Freeman, F. G. (1995). Sex differences in visuo-spatial ability: Task difficulty, speed accuracy trade-off, and other performance factors. *Canadian Journal of Experimental Psychology, 49,* 530–539.

Pujol, J., Deus, J., Losilla, J. M., & Capdevila, A. (1999). Cerebral lateralization of language in normal left-handed people studied by functional MRI. *Neurology, 52,* 1038-1043.

Puts, D. A., McDaniel, M. A., Jordan, C. L., & Breedlove, M. (2008). Spatial ability and prenatal androgens: Meta-analysis of congenital adrenal hyperplasia and digit ratio (2D:4D) studies. *Archives of Sexual Behaviour, 37,* 100-111.

Putz, D. A., Gaulin, S. J. C., Sporter, R. J., & McBurney, D.H. (2004) Sex hormones and finger length: What does 2D:4D indicate? *Evolution and Human Behavior, 25,* 182–199.

Rahman, Q., & Wilson, G. D. (2003). Sexual orientation and the 2nd to 4th finger length ratio: Evidence for organizing effects of sex hormones or developmental instability? *Psychoneuroendocrinology*, *28,* 288–303.

Rasmussen, T., & Milner, B. (1977). The role of early left-brain injury in determining lateralization of cerebral speech functions. *Annals of the New York Academy of Science, 30,* 355-369.

Resnick, S. M. (1983). Psychological functioning in individuals with congenital adrenal hyperplasia: Early hormonal influences on cognition and personality (Doctoral dissertation, University of Minnesota, 1982). *Dissertation Abstracts International, 44,* 642B.

Resnick, S. M., Berenbaum, S. A., Gottesman, I. I., & Bouchard, T. J., Jr. (1986). Early hormonal influences on cognitive functioning in congenital adrenal hyperplasia. *Developmental Psychology, 22*, 191-198.

Resnick, S. M., Gottesman, I. I. & McGue, M. (1993). Sensation seeking in opposite-sex twins: An effect of prenatal hormones? *Behaviour Genetics, 23*, 323-329.

Reusser, K. (2000). Success and failure in school mathematics: Effects of instruction and school environment. *European Child & Adolescent Psychiatry, 9,* S17-S26.

Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex, 15*, 1779-1790.

Robinson, S. J., & Manning, J. T. (2000). The ratio of 2nd to 4th digit lengths and male homosexuality. *Evolution and Human Behaviour, 21, 333*-345.

Rodeck, C. H., Gill, D., Rosenberg, D. A. & Collins, W. P. (1985). Testosterone levels in midtrimester and fetal plasma and amniotic fluid. *Prenatal Diagnosis, 5,* 175-181.

Rohde Parfet, K. A., Ganjam, V. K., Lamberson, W. R., Rieke, A. R., vom Saal, F. S. & Day, B. N. (1990). Intrauterine position effects in female swine: Subsequent reproductive performance, and social and sexual behaviour. *Applied Animal Behaviour Science, 26,* 349-362.

Roland, P. E., & Friberg, L. (1985). Localization of cortical areas activated by thinking. *Journal of Neurophysiology, 53,* 1219-1243.

Romeo, R. D. (2003). Puberty: A period of both organizational and activational effects on steroid hormones on neurobehavioral development. *Journal of Neuroendocrinology, 15*, 1185-1192.

Romeo, R. D. (2005). Neuroendocrine and behavioural development during puberty: A tale of two axes. *Vitamins and Hormones, 71,* 1-25.

Romeo, R. D., Richardson, H. N., & Sisk, C. L. (2002). Puberty and the maturation of the male brain and sexual behaviour: recasting a behavioural potential. *Neuroscience and biobehavioral reviews, 26,* 379-389.

Roof, R. L. & Havens, M. D. (1992). Testosterone improves maze performance and induces development of a male hippocampus in females. *Brain Research, 572*, 310-313.

Rose, R. M. (1985). Psychoneuroendocrinology. In J. D. Wilson & D. W. Foster (Eds.) *Williams textbook of endocrinology* (pp.653-681). Philadelphia: W. B. Saunders.

Rouselle, L., & Nőel, M-P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs. non-symbolic number magnitude processing. *Cognition, 102,* 361-395.

Rueckert, L., Lange, N., Partiot, A., Appollonio, I., Litvar, I., Le Bihan, D., & Grafman, J. (1996). Visualizing cortical activation during mental calculation with functional MRI. *NeuroImage, 3,* 97-103.

Ryan, B. C. & Vandenbergh, J. G. (2002). Intrauterine position effects. *Neuroscience and Biobehavioural Reviews, 26,* 665-678.

Sakai, L. M., Baker, L. A., Jacklin, C. N., & Shulman, I. (1992). Sex steroids at birth: Genetic and environmental variation and co variation. *Developmental Psychobiology, 24,* 559-570.

Sanders, G., Bereczkei, T., Csatho, A. & Manning, J. (2005). The ratio of the 2nd to 4th finger length predicts spatial ability in men but not women. *Cortex, 41,* 789-795.

Sanders, G., & Kadam, A. (2001). Prepubescent children show the adult relationship between dermatoglyphic asymmetry and performance on sexually dimorphic tasks. *Cortex, 37,* 91-100.

Sanders, G., & Waters, F. (2001). Fingerprint asymmetry predicts within sex differences in the performance of sexually dimorphic tasks. *Personality and Individual Differences, 31,* 1181-1191.

Santos, L., Hauser, M., & Spelke, E. (2001). Recognition and categorization of biologically significant objects by rhesus monkeys *(Macaca Mulatta)*: The domain of food. *Cognition, 82,* 127-155.

Sathian, K., Simon, T. J., Peterson, S., Patel, G. A., Hoffman, J. M., & Grafton, S. T. (1999). Neural evidence linking visual object enumeration and attention. *Journal of Cognitive Neuroscience, 11,* 36-51.

Saucier, D., Bowman, M., & Elias, L. (2003). Sex differences in the effect of articulatory or spatial dual – task interference during navigation. *Brain and Cognition, 53,* 346-350.

Sawamura, H., Shima, K., & Tanji, J. (2002). Numerical representation for action in the parietal cortex of the monkey. *Nature, 415,* 918-922.

Schindler, A. E. (1982). *Hormones in amniotic fluid.* Heidelberg: Springer-Verlag.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual object hood. *Cognitive Psychology. 38,* 259-290.

Segovia, S., Orensanz, L. M., Valencia, A. & Guillamon, A. (1984). Effects of sex steroids on the development of the accessory olfactory bulb in the rat: volumetric study. *Brain research, 16,* 312-314.

Sharon, T., & Wynn, K. (1998). Individuation of Actions from Continuous Motion. *Psychological Science, 9,* 357-362.

Sheridan, P. J. (1979). Androgen receptors in the brain: What are we measuring? *Endocrine Reviews, 4,* 171-178.

Shute, V., Pellegrino, J. W., Hubert, L., & Reynolds, R. W. (1983). The relationship between androgen levels and human spatial abilities. *Bulletin of the Psychonomic Society, 21,* 465-468.

Siegal, S., & Castellan, N. J. (1988). *Nonparametric Statistics for the Behavioural Sciences. Second Edition.* London: McGraw Hill.

Silverman, I., & Phillips, K. (1993). Effects of estrogen changes during the menstrual cycle on spatial performance. *Ethology and Sociobiology, 14,* 257-269.

Silverman, I. W., & Rose, A. P. (1980). Subitizing and counting skills in 3-year olds. *Developmental Psychology, 16,* 539-540.

Simerly, R. B., Chang, C., Muramatsu, M. & Swanson, L. W. (1990). Distribution of androgen and estrogen receptor mRNA-containing cells in the rat brain: an *in situ* hybridization study. *Journal of Comparative Neurology, 294,* 76-95.

Simerly, R. B., Swanson, L. W., Honda, R. J. & Gorski, R. A. (1985). Influence of perinatal androgen on the sexually dimorphic distribution of the tyrosine hydroxylase-immunoreactive cells and fibres in the anteroventral periventricular nucleus of the rat. *Neuroendocrinology, 40,* 501-510.

Simerly, R. B., Zee, M. C., Pendleton, J. W., Lubahn, D. B. & Korach, K. S. (1997). Estrogen receptor-dependent sexual differentiation of dopaminergic neurons in the preoptic region of the mouse. *Proceedings of the National Academy of Sciences, 94,* 14077-14082.

Simon, T. J., Hespos, S. J. & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development, 10,* 253-269.

Sinforiani, E., Livieri, C., Maufi, M., Bisio, P., Sibilla, L., Chiesa, L. & Martelli, A. (1994). Cognitive and neuroradiological findings in congenital adrenal hyperplasia. *Psychoneuroendocrinology, 19,* 55-64.

Sisk, C. L., Schulz, K. M. & Zehr, J. L. (2003). Puberty: A Finishing school for male social behavior. *Annals of the New York Academy of Sciences, 1007,* 189-198.

Sisk, C. L. & Zehr, J. L. (2006). Pubertal hormones organize the adolescent brain and behaviour. *Frontiers in Neuroendocrinology, 26*, 163-174.

Slabbekoorn, D., van Goozen, S.H.M, Sanders, G., Gooren, L.J.G. & Cohen-Kettenis, P.T. (2000). The dermatoglyphic characteristics of transsexuals: Is there evidence for an organizing effect of sex hormones? *Psychoneuroendocrinology, 25,* 365–375.

Smail, P. J., Reyes, F. I., Winter, J. S. D. & Faiman, C. (1981). The fetal hormone environment and its effect on the morphogenesis of the genital system. In S. J. Kogan & E. S. E. Hafez (Eds.) *Pediatric andrology* (pp. 9-19). The Hague, Netherlands: Martinus Nijhoff.

Sommer, I. E., Aleman, A., Bouma, A., & Kahn, R. S. (2004). Do women really have more bilateral language representation than men? A meta-analysis of functional imaging studies. *Brain, 127,* 1845-1852.

Sommer, I. E., Aleman, A., Somers, M., Boks, M. P., & Kahn, R. S. (2008). Sex differences in handedness, asymmetry of the Planum Temporale and functional language lateralization. *Brain Research, 1206,* 76-88.

Spelke, E. S., & Tsivkin, S. (2001). Language and number: a bilingual training study. *Cognition, 78,* 45-88.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4-28.

Springer, S. P., & Deutsch, G. (1997) *Left brain, right brain: Perspectives from cognitive neuroscience.* W. H. Freeman: San Francisco.

Stanescu-Cosson, R., Pinel, P., Van de Moortele, P. F. Le Bihan, D., Cohen, L., & Dehaene, S. (2000). Cerebral bases of calculation processes: Impact of number size on the cerebral circuits for exact and approximate calculation. *Brain, 123,* 2240-2255.

Starkey, P., & Cooper, R. G. Jr. (1980). Perception of numbers by human infants. *Science, 210,* 1033-1035.

Starkey, P., Spelke, E. S., & Gelman, R. (1983). Detection of intermodal numerical correspondences by human infants. *Science, 222,* 179-181.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36,* 97-127.

Steenhuis, R. E., & Bryden, M. P. (1989). Different dimensions of hand preferences that relate to skilled and unskilled activities. *Cortex, 25,* 289-304.

Stewart, J. & Cygan, D. (1980). Ovarian hormones act early in development to feminize adult open-field behavior in the rat. *Hormones and Behaviour, 14,* 20–32.

Stewart, J., Skvarenina, A. & Pottier, J. (1975). Effects of neonatal androgens on open-field behaviour and maze learning in the pre-pubescent and adult rat. *Physiology & Behaviour, 14,* 291-295.

Stones, I., Beckmann, M., & Stephens, L. (1982) Sex related differences in mathematical competencies of pre-calculus college students. *School Science and Mathematics, 82,* 295-299.

Strauss, M. S., & Curtis, L. E. (1981) Infant perception of numerosity. *Child Development* **52**: 1146-1152

Sulkowski G. M., & Hauser, M. D. (2001). Can rhesus monkeys spontaneously subtract?, *Cognition, 79,* 239–262.

Svartberg, J., Jorde, R., Sundsfjord, J., Bønaa, K. H. & Barrett-Connor, E. (2003). Seasonal Variation of Testosterone and Waist to Hip Ratio in Men: The Tromsø Study. *The Journal of Clinical Endocrinology & Metabolism, 88,* 3099–3104.

Svenson, O., & Sjöberg, K. (1983). Speeds of subitizing and counting processes in different age groups. *The Journal of Genetic Psychology, 142,* 203-211.

Swaab, D. F., Chung, W. C., Kruijver, F. P, & Hofman, M. A. (2002). Sexual differentiation of the human hypothalamus. *Advances in Experimental Medicine and Biology, 511,* 75-100.

Talmadge, C. L., Long, G. R., Murphy, W. J. & Tubis, A. (1993). New off-line method for detecting spontaneous otoacoustic emissions in human subjects. *Hearing Research, 71,* 170–182.

Tanner, J. M. (1990). *Foetus into man: Physical growth from conception to maturity.* Masseuses: Harvard University Press.

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-olds and adults. *Proceedings of the National Academy of Science USA, 95,* 7836-7841.

Teuber, H. L. (1974) Why two brains. In F. O. Schmitt and F. G. Worden (Eds.) *The Neuroscience: Third Study Program* MIT: Cambridge

Thioux, M., Pillon, A., Samson, D., De Partz, M. P., Noel, M. P., & Seron, X. (1998). The isolation of numerals at the semantic level. *NeuroCase, 4,* 371-389.

Trick, L., M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 331-351.

Trick, L. M., Pylyshyn, Z. W. (1994) Why are small and large numbers enumerated differently? A limited capacity preattentive stage in vision. *Psychological Review, 101,* 80-102.

Trivers, R., Manning, J., & Jacobson, A. (2006). A longitudinal study of digit ratio (2D:4D) and other finger ratios in Jamaican children. *Hormones and Behavior, 49,* 150-156.

Uller, C., Hauser, M., & Carey, S. (2001). Spontaneous representation of number in cotton-top tamarins (Saguinus oedipus). *Journal of Comparative Psychology, 11,* 248-257.

Valencia, A., Sergovia, S. & Guillamon, A. (1986). Effects of sex steroids on the development of the accessory olfactory bulb mitral cells in the rat. *Brain research, 24,* 287-290.

van Anders, S. M., & Hampson, E. (2005). Testing the prenatal androgen hypothesis: measuring digit ratios, sexual orientation, and spatial abilities in adults. *Hormones and Behavior, 47,* 92-98.

van Anders, S. M., Vernon, P. A., & Wilbur, C. J. (2006). Finger-length ratios show evidence of prenatal hormone-transfer between opposite-sex twins. *Hormones and Behavior, 49,* 315-319.

van de Beek, C., Thijssen, J. H. H., Cohen-Kettenis, P. T., van Goozen, S. H. M. & Buitelaar, J. K. (2004). Relationships between sex hormones assessed in amniotic fluid and maternal and umbilical cord serum: What is the best source of information to investigate the effects of fetal hormonal exposure? *Hormones and Behaviour, 46,* 663-669.

van Marle, K., & Scholl, B. J. (2003). Attentive tracking of objects vs. substances. *Psychological Science, 14,* 498-504.

Vermeulen, A., & Verdonck, L. (1976). Plasma androgen levels during the menstrual cycle. *American Journal of Obstetrics and Gynecology, 125,* 491-494.

van Loosbroek, E., & Smitsman, A. D. (1990). Visual perception of numerosity in infancy. *Developmental Psychology, 26,* 916–922.

vom Saal, F. (1989). Sexual differentiation in litter-bearing mammals: Influence of sex of adjacent foetuses in utero. *Journal of Animal Science, 67,* 1824-1840.

von Aster, M. G. (2001). Die Neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern (ZAREKI). Swets & Zeitlinger, Swets Test Services: Lisse, Frankfurt.

Voracek, M., & Dressler, S. G. (2007). Digit ratio (2D:4D) in twins: heritability estimates and evidence for a masculinised trait expression in women from opposite-sex pairs. *Psychological Reports, 100,* 115-126.

Voyer, D., Voyer, S. & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and considerations of critical variables. *Psychological Bulletin, 117,* 250-270.

Walker, E. F., Sabuwalla, Z. & Huot, R. (2004). Pubertal neuromaturation, stress sensitivity, and psychopathology. *Development and Psychopathology, 16,* 807-824.

Washburn, D., & Rumbaugh, D. M. (1991). Ordinal judgements of numerical symbols by macaques (Macaca mulatta). *Psychological Science, 2,* 190-230.

Weinberg, S. M., Scott, N. M., Neiswanger, K., & Marazita, M. L. (2005). Intraobserver error associated with measurements of the hand. *American Journal of Human Biology, 17*, 368-371.

Wender, K. F. & Rothkegel, R. (2000) Subitizing and its sub processes. *Psychological Research*, *64,* 81-92.

West, R. E., & Young, R., J. (2002). Do domestic dogs show any evidence of being able to count? *Animal Cognition, 5,* 1435-1448.

Whalen, J., Gallistel, C, R., & Gelman, R. (1999). Nonverbal Counting in Humans: The Psychophysics of Number Representation. *Psychological Science, 10,* 130-137.

Whitcomb, R. W. & Crowley, W. F. (1993). Male hypogonadotropic hypogonadism. *Endocrinology and Metabolism Clinics of North America, 22,* 125-163.

White, P. C. & Speiser, P. W. (2000). Congenital adrenal hyperplasia due to 21-hydroylase deficiency. *Endocrine Reviews, 21,* 245-291.

Williams, C. L., Barnett, A. M. & Meck, W. H. (1990). Organizational effects of early gonadal steroids on sexual differentiation in spatial memory. *Behavioural Neuroscience, 104,* 84-97.

Williams, C. L. & Meck, W. H. (1991). The organizational effects of gonadal steroids on sexually dimorphic spatial ability. *Psychoneuroendocrinology, 16,* 155-176.

Williams, J. H. G., Greenhalgh, K. D., Manning J. T. (2003). Second to fourth finger ratio and possible precursors of developmental psychopathology in pre-school children. *Early Human Development, 72,* 557-565.

Williams, T. J., Pepitone, M. E., Christensen, S. E., Cooke, B. M., A.D. Huberman, A. D., Breedlove, N., J., Breedlove, T. J., Jordan, C. L. (2000). Finger-length ratios and sexual orientation. *Nature, 404,* 455-456.

Willingham, W. W., & Cole, N. S. (1997). *Gender and fair assessment.* Mahwah, NJ: Erlbaum.

Witelson, S. F. (1991) Neural sexual mosaicism: Sexual differentiation of the human tempo-parietal region for functional asymmetry. *Psychoneuroendocrinology, 16*,131-154.

Witelson, S. F. & Nowakowski, R. S. (1991). Left out axons make men right: A hypothesis for the origin of handedness and functional asymmetry. *Neuropsychologia, 27*, 327-333.

Wisniewski, A. B. (1998). Sexually-dimorphic patterns of cortical asymmetry, and the role for sex steroid hormones in determining cortical patterns of lateralisation. *Psychoneuroendocrinology, 23*, 519-547.

Wisniewski, A. B., Migeon, C. J., Meyer-Bahlburg, H. F. L., Gearhart, J. P., Berkovitz, G. D., Brown, T. R. & Money, J. (2000). Complete androgen insensitivity syndrome: Long-term medical, surgical and psychosexual outcome. *The Journal of Clinical Endocrinology & Metabolism, 85,* 2664-2669.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358,* 749-750.

Wynn, K. (1992). Infants' Individuation and Enumeration of Actions. *Psychological Science, 7, 164-169.*

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition, 89,* B15-B25.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74,* B1-B11.

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8,* 88-101.

Ypsilanti, A., Ganou, M., Koidou, I. & Grouios, G. (2008). Digit ratio (2D:4D) in individuals with intellectual disability: Investigating the role of testosterone in the establishment of cerebral lateralisation. *Laterality: Asymmetries of Body, Brain and Cognition, 13,* 527-544.

Zakany, J., & Duboule, D. (1999). Hox genes and the making of sphincters. *Nature, 401*, 761-762.

Zorzi, M., Priftis, K., & Umiltà, C. (2002). Brain damage: Neglect disrupts the mental number line. *Nature, 417*, 138-139.

List of Appendices

Appendix 1

Additional statistical information relating Chapter 2

Section 2.3

Results from the Kolmogorov-Smirnov analysis to explore normality in data from children aged 5-7 years old.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RH2D4D	.072	41	.200*	.970	41	.349
LH2D4D	.093	41	.200*	.978	41	.599
Counting	.487	41	.000	.394	41	.000
Readnos	.399	41	.000	.509	41	.000
Writenos	.383	41	.000	.579	41	.000
MAov	.199	41	.000	.914	41	.004
NLov	.192	41	.001	.926	41	.011
NCov	.147	41	.026	.929	41	.014
Est	.178	41	.002	.866	41	.000
OVERALL	.138	41	.048	.938	41	.028

Results from the Kolmogorov-Smirnov analysis to explore normality in data from children aged 8-11 years old.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RH2D4D	.118	32	.200*	.958	32	.237
LH2D4D	.089	32	.200*	.941	32	.082
Counting	.455	32	.000	.589	32	.000
Readnos	.291	32	.000	.710	32	.000
Writenos	.370	32	.000	.522	32	.000
MAov	.143	32	.096	.909	32	.011
NLov	.207	32	.001	.714	32	.000
NCov	.238	32	.000	.867	32	.001
Est	.151	32	.063	.933	32	.049
OVERALL	.180	32	.010	.823	32	.000

Means and standard deviations for 2D:4D measures and scores on the numerical test battery in children aged 5-7 years old.

Measure	Overall N = 41	Males N = 19	Females N = 22
RH2D4D	0.978 (0.03)	0.978 (0.03)	0.979 (0.03)
LH2D4D	0.986 (0.04)	0.982 (0.04)	0.989 (0.04)
Counting	96.83 (9.34)	97.37 (6.53)	96.36 (11.36)
Reading numbers	91.46 (19.69)	95.26 (12.19)	88.18 (24.23)
Writing numbers	89.76 (19.56)	92.63 (13.68)	87.27 (23.54)
Mental arithmetic overall	68.54 (24.53)	71.58 (25.88)	65.91 (23.59)
Number line tasks overall	68.45 (14.11)	68.75 (13.66)	68.18 (14.8)
Number comparison overall	81.88 (13.66)	80.83 (15.26)	82.79 (12.4)
Estimation	23.66 (21.3)	28.95 (23.31)	19.09 (18.75)
Test battery overall	73.69 (11.13)	75.62 (12.48)	72.02 (9.8)

Means and standard deviations for 2D:4D measures and scores on the numerical test battery in children aged 8-11 years old.

Measure	Overall N = 32	Males N = 18	Females N = 14
RH2D4D	0.985 (0.03)	0.987 (0.03)	0.982 (0.04)
LH2D4D	0.979 (0.04)	0.984 (0.04)	0.973 (0.04)
Counting	97.03 (5.52)	95.28 (6.52)	99.29 (2.67)
Reading numbers	91.4 (10.57)	91.67 (10)	91.07 (11.63)
Writing numbers	95.63 (9.22)	96.94 (7.1)	93.93 (11.47)
Mental arithmetic overall	68.75 (18.47)	70.74 (19.15)	66.19 (17.92)
Number line tasks overall	79.46 (17.17)	77.78 (20.95)	81.61 (10.96)
Number comparison overall	91.96 (7.09)	91.47 (8.05)	92.6 (5.85)
Estimation	39.69 (22.36)	46.67 (24.25)	30.71 (16.39)
Test battery overall	83.32 (8.42)	84 (9.17)	82.46 (7.59)

Section 2.3.2

Full list of correlations (Spearman's rho (ρ)) between age and; overall performance on the numerical test battery, and performance on each numerical subcategory, significant correlations (also cited in-text) are indicated in bold.

Task	**5-7 years old (n = 41)**		**8-11 years old (n = 32)**	
	$\boldsymbol{\rho}$	**p**	$\boldsymbol{\rho}$	**p**
Counting	0.256	0.106	0.316	0.078
Reading numbers	0.071	0.659	0.258	0.154
Writing numbers	-0.047	0.769	**0.375**	**0.045**
Mental arithmetic overall	0.175	0.274	0.001	0.995
Number line tasks overall	-0.006	0.968	**0.419**	**0.017**
Number comparison overall	-0.066	0.68	0.132	0.471
Estimation	0.004	0.98	0.029	0.877
Test battery overall	0.142	0.377	0.311	0.083

Section 2.3.3

U, z, and p values relating to the effect of sex for Mann-Whitney *U* analysis of overall scores on the test battery and scores on each numerical subcategory in children aged 5-7 years. n = 41

Sub-Catergory	***U***	**Z**	**p**
Counting	205.5	-0.149	0.882
Reading numbers	182	-0.906	0.365
Writing numbers	187.5	-0.685	0.493
Mental arithmetic overall	179	-0.793	0.428
Number line tasks overall	195	-0.372	0.71
Number comparison overall	199.5	-0.252	0.801
Estimation	160	-1.317	0.188
Test battery overall	153.5	-1.453	0.146

U, z and p values relating to the effect of sex for Mann-Whitney *U* analysis of overall scores on the test battery and scores on each numerical subcategory in children aged 8-11 years, significant effects (also cited in-text) are indicated in bold. n = 32

Sub-Catergory	***U***	**Z**	**p**
Counting	**85**	**-2.059**	**0.039**
Reading numbers	124	-0.061	0.952
Writing numbers	125	-0.023	0.981
Mental arithmetic overall	100.5	-0.972	0.331
Number line tasks overall	123	-0.115	0.909
Number comparison overall	122.5	-0.136	0.892
Estimation	76	-1.92	0.055
Test battery overall	105	-0.799	0.424

Appendix 2

Additional statistical information relating Chapter 4

Section 4.2.3

Average 2D:4D values and t-test analysis of group differences in 2D:4D for high and low 2D:4D groups.

	Mean RH 2D:4D		Group comparisons		Mean LH 2D:4D		Group comparisons	
	Low	High	t	p	Low	High	t	p
Males (n = 30)	0.955	0.995	7.12	<0.001	0.959	1.003	7.93	<0.001
Females (n = 46)	0.966	1.015	8.76	<0.001	0.964	1.015	9.39	<0.001

Section 4.3

Means and standard deviations for 2D:4D measures, subitizing and counting reaction times and percentage error scores in the entire sample, i.e. including both males and females.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.985 (0.03)	0.962 (0.02)	1.007 (0.02)	0.967 (0.02)	1.004 (0.03)
Left hand 2D:4D	0.986 (0.03)	0.967 (0.02)	1.004 (0.03)	0.962 (0.01)	1.01 (0.02)
Subitizing RT – RVF	903.7 (150.31)	908.15 (170.8)	899.48 (130.01)	883.16 (112.62)	924.24 (179.56)
Subitizing RT – LVF	879.44 (144.42)	891.65 (167.81)	867.85 (119.17)	862.97 (105.58)	895.91 (174.85)
Subitizing RT – Overall	891.59 (145.58)	899.9 (167.88)	883.71 (122.48)	873.06 (107.2)	910.12 (175.4)
Counting RT – RVF	1575.27 (325.57)	1591.31 (312.13)	1560.05 (341.2)	1563.35 (280.22)	1587.18 (368.84)
Counting RT - LVF	1578.50 (336.34)	1579.12 (325.7)	1577.9 (350.38)	1540.09 (283.15)	1616.91 (382.22)
Counting RT – Overall	1576.93 (327.18)	1585.22 (315.49)	1569.06 (341.83)	1551.72 (278.18)	1602.14 (371.88)
Subitizing errors – RVF	3.36 (3.13)	3.42 (3.25)	3.3 (3.05)	3.22 (3.17)	3.51 (3.13)
Subitizing errors – LVF	2.6 (3.39)	2.46 (3.13)	2.74 (3.66)	2.4 (3.03)	2.81 (3.75)
Subitizing errors – Overall	2.98 (2.81)	2.94 (2.65)	3.02 (2.99)	2.81 (2.58)	3.16 (3.05)
Counting errors – RVF	6.93 (7.05)	6.13 (5.8)	7.69 (8.06)	5.96 (6.01)	7.89 (7.92)
Counting errors – LVF	6.75 (7.44)	5.5 (5.62)	7.95 (8.74)	5.88 (5.5)	7.63 (8.97)
Counting errors – Overall	6.84 (6.6)	5.81 (4.99)	7.82 (7.76)	5.92 (5.02)	7.76 (7.83)

Means and standard deviations for 2D:4D measures, subitizing and counting reaction times and percentage error scores in males only.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.977 (0.03)	0.955 (0.01)	0.995 (0.02)	0.962 (0.02)	0.991 (0.02)
Left hand 2D:4D	0.981 (0.03)	0.963 (0.02)	0.996 (0.02)	0.959 (0.02)	1.003 (0.01)
Subitizing RT – RVF	911.88 (202.47)	902.97 (243.41)	919.68 (166.64)	834.5 (96.03)	989.27 (250.73)
Subitizing RT – LVF	881.1 (195.31)	894.23 (249.13)	869.62 (140.22)	813.86 (98.92)	948.35 (244.01)
Subitizing RT – Overall	896.5 (197.28)	898.6 (245.73)	894.67 (151.34)	824.18 (96.91)	968.83 (245.1)
Counting RT – RVF	1497.39 (363.8)	1506.5 (411.81)	1489.42 (329.79)	1419.23 (257.52)	1575.54 (441.3)
Counting RT - LVF	1509.54 (386.05)	1488.2 (425.65)	1528.2 (360.99)	1403.52 (281.67)	1615.55 (453.1)
Counting RT – Overall	1503.58 (371.71)	1497.35 (416.25)	1509.03 (341.91)	1411.38 (266.76)	1595.78 (443.66)
Subitizing errors – RVF	3.78 (3.51)	4.6 (3.43)	3.06 (3.53)	4.59 (3.41)	2.96 (3.53)
Subitizing errors – LVF	2.44 (3.37)	2.38 (3.08)	2.5 (3.71)	2.37 (2.97)	2.52 (3.84)
Subitizing errors – Overall	3.11 (3.11)	3.49 (3.01)	2.78 (3.25)	3.48 (2.84)	2.74 (3.41)
Counting errors – RVF	7 (9.03)	6.43 (6.98)	7.5 (10.72)	5.56 (7.09)	8.44 (10.68)
Counting errors – LVF	7.56 (7.83)	8.1 (7.13)	7.08 (8.6)	7.11 (7.11)	8 (8.71)
Counting errors – Overall	7.28 (7.8)	7.26 (6.29)	7.29 (9.13)	6.33 (6.37)	8.22 (9.14)

Means and standard deviations for 2D:4D measures, subitizing and counting reaction times and percentage error scores in females only.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.991 (0.03)	0.966 (0.02)	1.015 (0.02)	0.97 (0.02)	1.012 (0.03)
Left hand 2D:4D	0.989 (0.03)	0.969 (0.02)	1.01 (0.03)	0.964 (0.01)	1.015 (0.02)
Subitizing RT – RVF	898.36 (105.64)	0.911 (112.69)	885.42 (98.88)	914.89 (113.12)	881.83 (97.25)
Subitizing RT – LVF	878.35 (100.87)	890.09 (96.96)	866.62 (105.46)	895 (99.01)	861.71 (102.13)
Subitizing RT – Overall	888.39 (101.08)	900.7 (102.15)	876.08 (100.74)	904.95 (103.42)	871.83 (98.13)
Counting RT – RVF	1626.06 (291.01)	1642.93 (227.7)	1609.18 (347.54)	1657.34 (257.8)	1594.77 (323.6)
Counting RT - LVF	1623.47 (295.39)	1634.46 (240.97)	1612.48 (346.64)	1629.15 (251.68)	1617.79 (339.21)
Counting RT – Overall	1624.77 (288.86)	1638.7 (229.32)	1610.83 (343.05)	1643.25 (250.35)	1606.28 (327.55)
Subitizing errors – RVF	3.09 (2.87)	2.71 (2.99)	3.48 (2.75)	2.32 (2.72)	3.86 (2.86)
Subitizing errors – LVF	2.71 (3.44)	2.51 (3.23)	2.9 (3.7)	2.42 (3.14)	3 (3.77)
Subitizing errors – Overall	2.9 (2.63)	2.61 (2.41)	3.19 (2.87)	2.37 (2.35)	3.43 (2.84)
Counting errors – RVF	6.88 (5.51)	5.94 (5.12)	7.83 (5.83)	6.23 (5.35)	7.54 (5.7)
Counting errors – LVF	6.23 (7.22)	3.91 (3.85)	8.55 (8.97)	5.07 (4.13)	7.39 (9.32)
Counting errors – Overall	6.56 (5.76)	4.93 (3.91)	8.19 (6.85)	5.65 (4.04)	7.46 (7.05)

Results from the Kolmogorov-Smirnov analysis to explore normality.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Final right hand 2D:4D	.083	76	.200*	.978	76	.207
Final left hand 2D:4D	.089	76	.200*	.977	76	.179
RT 2 overall	.126	76	.005	.848	76	.000
RT 3 overall	.107	76	.031	.816	76	.000
RT 4 overall	.177	76	.000	.833	76	.000
RT subitizing right	.134	76	.002	.846	76	.000
RT subitizing left	.119	76	.010	.847	76	.000
RT subitizing overall	.119	76	.010	.846	76	.000
RT 6 overall	.086	76	.200*	.964	76	.028
RT 7 overall	.061	76	.200*	.982	76	.374
RT 8 overall	.043	76	.200*	.968	76	.049
RT count right	.066	76	.200*	.982	76	.357
RT count left	.082	76	.200*	.980	76	.262
RT count overall	.084	76	.200*	.979	76	.246
Error 2 overall	.317	76	.000	.732	76	.000
Error 3 overall	.270	76	.000	.779	76	.000
Error 4 overall	.281	76	.000	.745	76	.000
Error subitizing right	.195	76	.000	.857	76	.000
Error subitizing left	.308	76	.000	.741	76	.000
Error subitizing overall	.199	76	.000	.829	76	.000
Error 6 overall	.245	76	.000	.766	76	.000
Error 7 overall	.193	76	.000	.795	76	.000
Error 8 overall	.244	76	.000	.703	76	.000
Error count right	.210	76	.000	.793	76	.000
Error count left	.236	76	.000	.764	76	.000
Error count overall	.175	76	.000	.769	76	.000

*. This is a lower bound of the true significance.

Section 4.3.3

Non-reported F, p, MSe, effect size (partial eta squared - ηp^2), and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), sex, process (subitizing vs. counting) and visual field (left vs. right) on reaction times.

Effect	**F value**	**df**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect of visual field	2.835	1,72	0.097	0.003	0.038	0.383
Main effect sex	1.144	1,72	0.288	0.204	0.016	0.184
Visual field x sex	0.019	1,72	0.89	0.003	0.0003	0.052
Process x visual field x sex interaction	0.691	1,72	0.409	0.003	0.01	0.13

Unreported t and p values from the Post Hoc t-test conducted to evaluate the significant interaction between process (subitizing vs. counting) and sex on reaction times.

Comparison	**t value**	**df**	**p**
Sex differences in subitizing reaction times	0.236	74	0.814
Sex differences in counting reaction times	1.595	74	0.115

F, p, MSe, effect size (partial eta squared - ηp^2), and power (1-β) values from the ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), sex, process (subitizing vs. counting) and visual field (left vs. right) on reaction times, significant effect indicated in bold.

Effect	**F value**	**df**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect 2D:4D	1.582	1,72	0.213	0.192	0.05	0.237
Main effect Process	**590.167**	**1,72**	**<0.001**	**0.056**	**0.891**	**1**
Main effect of Visual field	2.899	1,72	0.093	0.003	0.039	0.39
Main effect Sex	1.211	1,72	0.275	0.192	0.017	0.192
2D:4D x Process interaction	0.104	1,72	0.747	0.056	0.001	0.062
2D:4D x Visual field interaction	3.184	1,72	0.079	0.003	0.042	0.421
2D:4D x Sex interaction	3.763	1,72	0.056	0.192	0.05	0.482
Visual field x Sex interaction	0.027	1,72	0.87	0.003	0.0004	0.053
Visual field x Process interaction	**5.114**	**1,72**	**0.027**	**0.003**	**0.066**	**0.607**
Process x Sex interaction	**5.479**	**1,72**	**0.022**	**0.056**	**0.071**	**0.637**
2D:4D x Process x Visual field	**5.699**	**1,72**	**0.02**	**0.003**	**0.073**	**0.654**
Process x sex x 2D:4D interaction	0.153	1,72	0.696	0.056	0.002	0.067
Visual field x sex x 2D:4D interaction	0.103	1,72	0.749	0.003	0.001	0.062
Process x visual field x sex interaction	0.913	1,72	0.342	0.003	0.013	0.156
Process x visual field x sex x 2D:4D interaction	0.212	1,72	0.647	0.003	0.003	0.074

Appendix 3

Additional statistical information relating Chapter 5

Section 5.2.3

Average 2D:4D values and t-test analysis of group differences in 2D:4D for high and low 2D:4D groups.

	Mean RH 2D:4D		Group comparisons		Mean LH 2D:4D		Group comparisons	
	Low	High	t	p	Low	High	t	p
Males (n = 34)	0.944	0.988	6.244	<0.001	0.931	0.981	7.792	<0.001
Females (n = 36)	0.96	1	7.533	<0.001	0.949	0.99	8.053	<0.001

Section 5.3.1

Means and standard deviations for 2D:4D measures, subitizing and control task reaction times and percentage error scores in the entire sample, i.e. including both males and females.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.973 (0.03)	0.953 (0.02)	0.994 (0.02)	0.96 (0.03)	0.986 (0.02)
Left hand 2D:4D	0.964 (0.03)	0.954 (0.03)	0.974 (0.03)	0.941 (0.02)	0.986 (0.02)
Subitizing RT – Left hand	761.09 (74.09)	767.03 (86.72)	755.14 (59.57)	771.04 (83.79)	751.68 (63.37)
Subitizing RT – Right hand	730.23 (74.87)	731.8 (89.83)	728.67 (57.46)	745.93 (85.74)	715.41 (60.46)
Subitizing RT – Overall	745.66 (71.33)	749.42 (85.71)	741.9 (54.3)	758.49 (81.31)	733.55 (59.04)
Control RT – Left hand	550.64 (72.46)	565.18 (79.97)	536.1 (61.84)	568.3 (74.87)	533.96 (66.9)
Control RT – Right hand	533.09 (79.2)	538.51 (88.9)	527.67 (69.03)	537.43 (85.23)	528.99 (74.04)
Control RT – Overall	541.07 (69.83)	551.85 (78.57)	531.89 (59.31)	552.87 (73.54)	531.47 (65.47)
Subitizing errors – Left hand	4.02 (3.27)	4.11 (3.27)	3.92 (3.31)	4 (2.9)	4.03 (3.63)
Subitizing errors – Right hand	4.11 (3.59)	4.5 (3.51)	3.73 (3.68)	4.67 (3.51)	3.59 (3.65)
Subitizing errors – Overall	4.07 (2.98)	4.3 (2.81)	3.83 (3.16)	4.33 (2.66)	3.81 (3.27)
Control errors – Left hand	2.52 (3.38)	2.29 (3.11)	2.76 (3.66)	2.75 (3.53)	2.31 (3.27)
Control errors – Right hand	2.76 (3.54)	2.67 (3.69)	2.86 (3.44)	2.94 (3.36)	2.59 (3.75)
Control errors – Overall	2.64 (2.62)	2.48 (2.37)	2.81 (2.88)	2.84 (2.68)	2.45 (2.6)

Means and standard deviations for 2D:4D measures, subitizing and control task reaction times and percentage error scores in males only.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.966 (0.03)	0.944 (0.02)	0.988 (0.02)	0.949 (0.02)	0.981 (0.03)
Left hand 2D:4D	0.958 (0.03)	0.944 (0.02)	0.971 (0.03)	0.931 (0.02)	0.981 (0.02)
Subitizing RT – Left hand	758 (84.83)	760.74 (104.59)	755.25 (62.35)	768.11 (110.04)	749.01 (55.71)
Subitizing RT – Right hand	734.39 (89.03)	746.45 (114.6)	722.34 (53.9)	751.22 (112.38)	719.43 (61.15)
Subitizing RT – Overall	746.2 (84.37)	753.61 (107.41)	738.8 (55.03)	759.67 (108.49)	734.23 (55.75)
Control RT – Left hand	554 (78.28)	573.08 (93.52)	534.93 (55.84)	571.75 (88.67)	538.23 (66.3)
Control RT – Right hand	539.12 (93.24)	533.79 (113.64)	544.45 (70.39)	537.17 (108.84)	540.85 (80.1)
Control RT – Overall	546.56 (79.12)	553.43 (96.66)	539.69 (58.9)	554.45 (90.29)	539.54 (69.63)
Subitizing errors – Left hand	3.68 (3.32)	3.84 (3.23)	3.53 (3.49)	3.33 (2.58)	4 (3.91)
Subitizing errors – Right hand	3.73 (3.2)	4.24 (3.34)	3.22 (3.06)	4.33 (3.17)	3.19 (3.21)
Subitizing errors – Overall	3.71 (2.9)	4.04 (2.88)	3.37 (2.98)	3.83 (2.3)	3.59 (3.42)
Control errors – Left hand	2.06 (2.96)	1.96 (2.65)	2.16 (3.32)	2.5 (3.1)	1.67 (2.86)
Control errors – Right hand	2.75 (3.71)	3.73 (4.55)	1.76 (2.39)	2.92 (3.63)	2.59 (3.89)
Control errors – Overall	2.4 (2.18)	2.84 (2.19)	1.96 (2.14)	2.71 (2.18)	2.13 (2.2)

Means and standard deviations for 2D:4D measures, subitizing and control task reaction times and percentage error scores in females only.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.98 (0.03)	0.96 (0.2)	1 (0.02)	0.97 (0.02)	0.99 (0.02)
Left hand 2D:4D	0.97 (0.03)	0.963 (0.03)	0.976 (0.02)	0.949 (0.02)	0.99 (0.01)
Subitizing RT – Left hand	764 (63.4)	772.96 (68.32)	755.04 (58.64)	773.64 (54.12)	754.36 (71.76))
Subitizing RT – Right hand	726.3 (59.55)	717.97 (57.97)	734.64 (61.58)	741.22 (55.47)	711.39 (61.26)
Subitizing RT – Overall	745.15 (57.6)	745.47 (61.66)	744.84 (55.03)	757.44 (49.45)	732.87 (63.76)
Control RT – Left hand	547.47 (67.46)	557.73 (66.61)	537.21 (68.63)	565.24 (62.62)	529.7 (69.14)
Control RT – Right hand	527.4 (64.04)	542.98 (60.08)	511.82 (65.73)	537.67 (60.4)	517.13 (67.62)
Control RT – Overall	537.44 (60.58)	550.36 (59.58)	524.52 (60.44)	551.47 (57.46)	523.41 (61.95)
Subitizing errors – Left hand	4.33 (3.24)	4.37 (3.39)	4.3 (3.19)	4.59 (3.11)	4.07 (3.44)
Subitizing errors – Right hand	4.48 (3.94)	4.74 (3.75)	4.22 (4.22)	4.96 (3.85)	4 (4.09)
Subitizing errors – Overall	4.41 (3.05)	4.56 (2.8)	4.26 (3.35)	4.78 (2.93)	4.04 (3.2)
Control errors – Left hand	2.96 (3.72)	2.59 (3.53)	3.33 (3.96)	2.96 (3.94)	2.96 (3.6)
Control errors – Right hand	2.78 (3.43)	1.67 (2.36)	3.89 (4)	2.96 (3.21)	2.59 (3.71)
Control errors – Overall	2.87 (3)	2.13 (2.54)	3.61 (3.3)	2.96 (3.11)	2.78 (2.97)

Results from the Kolmogorov-Smirnov analysis to explore normality.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
LH2d4d	.071	70	.200*	.976	70	.206
RH2d4d	.065	70	.200*	.980	70	.324
rt2both	.072	70	.200*	.977	70	.221
rt3both	.087	70	.200*	.972	70	.120
rt4both	.095	70	.192	.963	70	.037
rtsubleft	.072	70	.200*	.985	70	.592
rtsubright	.090	70	.200*	.953	70	.010
rtsubboth	.100	70	.082	.963	70	.038
redrtboth	.067	70	.200*	.987	70	.704
yelrtboth	.063	70	.200*	.973	70	.140
blurtboth	.055	70	.200*	.981	70	.380
colrtleft	.097	70	.172	.979	70	.301
colrtright	.056	70	.200*	.987	70	.685
colrtboth	.080	70	.200*	.986	70	.628
err2both	.211	70	.000	.891	70	.000
err3both	.218	70	.000	.844	70	.000
err4both	.204	70	.000	.800	70	.000
errssubleft	.160	70	.000	.920	70	.000
errssubright	.185	70	.000	.885	70	.000
errssubboth	.128	70	.007	.935	70	.001
rederrboth	.314	70	.000	.738	70	.000
yelerrboth	.357	70	.000	.713	70	.000
bluerrboth	.397	70	.000	.644	70	.000
colerrsleft	.315	70	.000	.749	70	.000
colerrsright	.268	70	.000	.752	70	.000
colerrsboth	.200	70	.000	.865	70	.000

*. This is a lower bound of the true significance.

Non-significant t and p values from the t-test analysis conducted in order to explore any differences in reaction time between stimuli in the subitizing and control tasks.

Comparison	**t**	**df**	**p**
2 vs. 3 dots reaction times	0.193	69	0.848
Red vs. Yellow reaction times	2.279	69	0.026

Non-significant Z and p values from the Wilcoxn signed ranks analysis conducted in order to explore any differences percentage error scores between stimuli in the subitizing and control tasks.

Comparison	**Z**	**p**
2 vs. 3 dots percentage error	-0.418	0.676
3 vs. 4 dots percentage error	-0.873	0.383
Red vs. Yellow percentage error	-1.472	0.141
Yellow vs. Blue percentage error	-1.078	0.281

Spearman's correlation (ρ) analysis of speed/accuracy associations for performance on the colour recognition and subitizing task (n = 70).

Analysis	ρ	**p**
Subitizing left visual field	0.01	0.935
Subitizing right visual field	0.185	0.126
Subitizing overall	0.099	0.416
Colour left visual field	-0.094	0.437
Colour right visual field	-0.194	0.107
Colour overall	-0.206	0.087

Section 5.3.2

Non-reported F and p values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (subitizing vs. control), sex (males vs. females) and response hand (left vs. right) on reaction times.

Effect	**F value**	**df**	**p**	**MSe**	ηp^2	**1-β**
Main effect Sex	0.098	1,66	0.755	0.018	0.001	0.061
Task x Sex interaction	0.56	1,66	0.457	0.002	0.008	0.114
Task x Response hand interaction	2.774	1,66	0.101	0.001	0.04	0.375
Sex x Response hand interaction	1.141	1,66	0.289	0.001	0.017	0.183
Task x Sex x Response hand interaction	0.317	1,66	0.575	0.001	0.005	0.086

Subsequent analysis of the 3-way interaction between 2D:4D group, sex and response hand on control task reaction times. Result of two separate 2-way ANOVAs conducted to investigate any possible main effect of 2D:4D and response hand and their interactions on control response times in males and females considered independently, significant results indicated in bold. *Post hoc analysis of this interaction are presented below.

Analysis	**Effect**	**F value**	**df**	**p**	**MSe**	**ηp^2**	**1-β**
Males	Main effect of 2D:4D	0.251	1,32	0.620	0.013	0.008	0.077
	Main effect of response hand	1.833	1,32	0.185	0.002	0.054	0.26
	2D:4D x Response hand interaction	**4.927**	**1,32**	**0.034***	**0.002**	**0.133**	**0.577**
Females	Main effect of 2D:4D	1.669	1,34	0.205	0.007	0.047	0.241
	Main effect of response hand	**5.424**	**1,34**	**0.026**	**0.001**	**0.138**	**0.619**
	2D:4D x Response hand interaction	0.381	1,34	0.541	0.001	0.011	0.092

Results of the post hoc analysis (bonferroni corrected t-tests, $\alpha = 0.0125$) to further investigate the interaction between 2D:4D group and response hand in males for reaction times on the control task.

Comparison	**t**	**df**	**p**
2D:4D group differences in right hand response times	0.329	32	0.744
2D:4D group differences in left hand response times	1.444	32	0.158
Response hand differences in low 2D:4D males	2.1	16	0.052
Response hand differences in high 2D:4D males	0.824	16	0.422

Non-reported F and p values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (subitizing vs. control), sex (males vs. females) and response hand (left vs. right) on reaction times.

Effect	**F value**	**df**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Sex	0.125	1,66	0.725	0.018	0.002	0.064
Task x Sex interaction	0.499	1,66	0.482	0.002	0.008	0.107
Task x Response hand interaction	2.236	1,66	0.14	0.001	0.033	0.314
Sex x Response hand interaction	1.034	1,66	0.313	0.001	0.015	0.171
Task x Sex x Response hand interaction	0.389	1,66	0.535	0.001	0.006	0.094

Appendix 4

Additional statistical information relating Chapter 6

Section 6.3.4

Average 2D:4D values and t-test analysis of group differences in 2D:4D for high and low 2D:4D groups for each task data set.

		Mean RH 2D:4D		Group comparisons		Mean LH 2D:4D		Group comparisons	
		Low	High	t	p	Low	High	t	p
Subitizing	Males (n = 33)	0.938	0.982	8.285	<0.001	0.949	0.985	8.601	<0.001
	Females (n = 30)	0.948	0.999	5.138	<0.001	0.959	1.002	5.079	<0.001
Counting	Males (n = 32)	0.941	0.986	8.891	<0.001	0.949	0.986	8.606	<0.001
	Females (n = 29)	0.953	0.998	4.845	<0.001	0.959	1.002	5.042	<0.001
Number Comparison	Males (n = 33)	0.944	0.985	8.507	<0.001	0.951	0.986	8.719	<0.001
	Females (n = 33)	0.95	1.004	6.633	<0.001	0.958	1.005	6.219	<0.001
SNARC	Males (n = 35)	0.939	0.984	8.837	<0.001	0.95	0.986	8.998	<0.001
	Females (n = 32)	0.954	1.003	5.556	<0.001	0.96	1.002	5.326	<0.001

Section 6.4

Subitizing

Means and standard deviations for 2D:4D measures, subitizing and control task reaction times and percentage error scores in the entire sample, i.e. including both males and females.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.968 (0.03)	0.942 (0.02)	0.991 (0.02)	0.953 (0.03)	0.983 (0.03)
Left hand 2D:4D	0.9742 (0.03)	0.959 (0.02)	0.988 (0.03)	0.954 (0.01)	0.994 (0.02)
Subitizing RT – LVF	546.88 (102.57)	555 (109.12)	539.5 (97.33)	545.47 (92.98)	548.25 (112.56)
Subitizing RT – RVF	552 (105.47)	557.3 (118.63)	547.18 (93.55)	558.45 (106.99)	545.75 (105.31)
Subitizing RT – Overall	548.39 (102.13)	554.32 (112.8)	543 (92.83)	550.24 (97.05)	546.59 (108.35)
Control RT – LVF	459.17 (83.4)	452.08 (74.06)	465.61 (91.74)	446.11 (82.18)	471.81 (83.91)
Control RT – RVF	454.17 (81.14)	446.05 (74.86)	461.56 (86.94)	437.77 (79.7)	470.06 (80.57)
Control RT – Overall	456.83 (82.19)	448.88 (73.67)	464.05 (89.76)	442.4 (80.94)	470.8 (82.22)
Subitizing errors – LVF	9.3 (7.35)	9.64 (8.4)	9. (6.36)	11.6 (8.53)	7.08 (5.22)
Subitizing errors – RVF	9.82 (8.92)	11.07 (11.34)	8.68 (5.92)	11.71 (11.49)	7.98 (4.95)
Subitizing errors – Overall	9.58 (7.38)	10.36 (9.26)	8.88 (5.16)	11.67 (9.17)	7.56 (4.35)
Control errors – LVF	0.57 (1.01)	0.56 (0.94)	0.57 (1.08)	0.54 (0.93)	0.59 (1.09)
Control errors – RVF	0.54 (1.22)	0.35 (0.8)	0.71 (1.49)	0.48 (1.05)	0.6 (1.38)
Control errors – Overall	0.55 (0.85)	0.46 (0.6)	0.64 (1.03)	0.51 (0.66)	0.59 (1.02)

Means and standard deviations for 2D:4D measures, subitizing and control task reaction times and percentage error scores in males only.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.96 (0.03)	0.938 (0.02)	0.982 (0.01)	0.95 (0.02)	0.9705 (0.03)
Left hand 2D:4D	0.967 (0.02)	0.956 (0.02)	0.978 (0.02)	.949 (0.01)	0.985 (0.01)
Subitizing RT – LVF	569.56 (101.58)	583.09 (115.22)	556.82 (88.52)	565 (92.2)	573.85 (112.37)
Subitizing RT – RVF	577.74 (107.9)	582.56 (125.05)	573.21 (92.61)	582.47 (108.51)	573.29 (110.46)
Subitizing RT – Overall	572.98 (103.23)	582.63 (120.89)	563.91 (86.19)	572.84 (98.89)	573.12 (110.19)
Control RT – LVF	484.77 (85.35)	472.81 (82.07)	496.03 (89.32)	473.59 (96)	495.29 (75.42)
Control RT – RVF	477.82 (86.09)	466.03 (86.97)	488.91 (86.37)	464.44 (94.32)	490.41 (78.32)
Control RT – Overall	481.67 (85.45)	469.22 (83.27)	493.38 (88.33)	469.69 (94.9)	492.94 (76.7)
Subitizing errors – LVF	9.63 (7.97)	10.19 (9.63)	9.09 (6.28)	12.38 (9.34)	7.04 (5.52)
Subitizing errors – RVF	10.87 (11.04)	12.43 (14.37)	9.4 (6.74)	13.33 (14.54)	8.55 (5.8)
Subitizing errors – Overall	10.26 (8.91)	11.33 (11.57)	9.25 (5.57)	12.9 (11.34)	7.78 (4.96)
Control errors – LVF	0.63 (1.1)	0.65 (1)	0.61 (1.22)	0.52 (0.94)	0.74 (1.26)
Control errors – RVF	0.32 (0.78)	0.27 (0.74)	0.37 (0.82)	0.27 (0.74)	0.37 (0.82)
Control errors – Overall	0.48 (0.74)	0.46 (0.66)	0.49 (0.83)	0.4 (0.53)	0.55 (0.91)

Means and standard deviations for 2D:4D measures, subitizing and control task reaction times and percentage error scores in females only.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.977 (0.04)	0.948 (0.02)	1.002 (0.03)	0.957 (0.03)	0.997 (0.03)
Left hand 2D:4D	0.982 (0.03)	0.964 (0.02)	0.998 (0.03)	0.959 (0.02)	1.005 (0.03)
Subitizing RT – LVF	521.93 (99.38)	522.89 (95.73)	521.09 (105.59)	524.63 (92.28)	519.23 (109.21)
Subitizing RT – RVF	523.68 (96.72)	528.43 (108)	519.53 (89.09)	532.83 (102.73)	514.53 (92.98)
Subitizing RT – Overall	521.33 (95.38)	521.96 (96.99)	520.78 (97.14)	526.13 (92.18)	516.53 (101.48)
Control RT – LVF	431 (72.57)	428.39 (57.73)	433.28 (85.34)	416.8 (53.05)	445.2 (87.54)
Control RT – RVF	428.17 (67.54)	423.21 (52.15)	432.5 (80.12)	409.33 (48.98)	447 (79.31)
Control RT – Overall	429.5 (70.07)	425.64 (54.92)	432.88 (82.78)	413.3 (51.31)	445.7 (83.52)
Subitizing errors – LVF	8.95 (6.72)	9 (7.05)	8.91 (6.64)	10.76 (7.81)	7.14 (5.04)
Subitizing errors – RVF	8.66 (5.76)	9.51 (6.59)	7.91 (5.01)	9.98 (7.07)	7.33 (3.85)
Subitizing errors – Overall	8.84 (5.27)	9.25 (5.87)	8.48 (4.84)	10.37 (6.22)	7.3 (3.7)
Control errors – LVF	0.49 (0.91)	0.45 (0.89)	0.53 (0.95)	0.56 (0.96)	0.43 (0.88)
Control errors – RVF	0.77 (1.55)	0.45 (0.89)	1.07 (1.94)	0.69 (1.29)	0.86 (1.82)
Control errors – Overall	0.63 (0.96)	0.45 (0.54)	0.8 (1.22)	0.63 (0.77)	0.64 (1.15)

Counting

Means and standard deviations for 2D:4D measures, counting and control task reaction times and percentage error scores in the entire sample, i.e. including both males and females.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.97 (0.03)	0.947 (0.02)	0.993 (0.02)	0.955 (0.02)	0.985 (0.03)
Left hand 2D:4D	0.975 (0.03)	0.96 (0.02)	0.989 (0.03)	0.953 (0.01)	0.995 (0.02)
Counting RT – LVF	1085.01 (286.19)	1113.21 (297.25)	1057.72 (277.18)	1130.12 (316.62)	1041.35 (250.79)
Counting RT – RVF	1103.2 (290.79)	1148.78 (308.98)	1059.1 (269.72)	1164.24 (326.98)	1044.14 (241.74)
Counting RT – Overall	1090.24 (289.52)	1128.03 (302.59)	1053.67 (276.26)	1141.14 (321.88)	1040.98 (249.78)
Control RT – LVF	567.83 (91.08)	563.13 (92.64)	572.37 (90.84)	546.08 (67.9)	588.87 (105.84)
Control RT – RVF	563.43 (90.53)	559.82 (94.02)	566.92 (88.44)	541.7 (72.63)	584.45 (101.81)
Control RT – Overall	566.06 (89.930	561 (93.06)	570.95 (88.06)	544.47 (69.88)	586.95 (102.66)
Counting errors – LVF	14.6 (12.67)	11.13 (8.46)	17.96 (15.1)	14.95 (12.55)	14.26 (12.98)
Counting errors – RVF	16.95 (13.36)	14.2 (11.59)	19.61 (14.57)	17 (14.27)	16.9 (12.64)
Counting errors – Overall	15.76 (12.29)	12.6 (9.12)	18.82 (14.21)	15.91 (12.47)	15.62 (12.32)
Control errors – LVF	0.41 (0.92)	0.56 (1.09)	0.28 (0.73)	0.42 (0.85)	0.41 (1)
Control errors – RVF	0.41 (1)	0.35 (0.96)	0.48 (1.05)	0.28 (0.73)	0.54 (1.21)
Control errors – Overall	0.41 (0.73)	0.45 (0.71)	0.38 (0.75)	0.35 (0.64)	0.48 (0.81)

Means and standard deviations for 2D:4D measures, subitizing and control task reaction times and percentage error scores in males only.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.963 (0.03)	0.941 (0.02)	0.986 (0.01)	0.948 (0.02)	0.978 (0.02)
Left hand 2D:4D	0.967 (0.02)	0.956 (0.02)	0.979 (0.02)	0.949 (0.01)	0.986 (0.01)
Counting RT – LVF	1063.43 (315.81)	1066.97 (299.3)	1059.89 (341.34)	1056.69 (322.67)	1070.17 (319.23)
Counting RT – RVF	1076.42 (304.1)	1100.08 (289.97)	1052.77 (325.33)	1102.58 (315.35)	1050.27 (300.35)
Counting RT – Overall	1071.73 (314.12)	1083.91 (295.65)	1059.55 (340.88)	1077.14 (322.44)	1066.31 (316.06)
Control RT – LVF	578.5 (92.62)	574.66 (103.44)	582.34 (83.65)	546.34 (56.26)	610.66 (111.17)
Control RT – RVF	573.75 (94.3)	571.25 (104.3)	576.25 (86.52)	537.59 (66.1)	609.91 (105.92)
Control RT – Overall	576.53 (92.62)	571.78 (104.12)	581.28 (82.71)	542.66 (61.24)	610.41 (107.38)
Counting errors – LVF	15.5 (12.45)	12.86 (9.33)	18.14 (14.78)	16.34 (10.79)	14.66 (14.23)
Counting errors – RVF	18.13 (14.18)	15.33 (10.84)	20.92 (16.77)	20.33 (13.61)	15.92 (14.82)
Counting errors – Overall	16.81 (12.84)	14.05 (9.31)	19.56 (15.42)	18.28 (11.51)	15.33 (14.27)
Control errors – LVF	0.33 (0.93)	0.65 (1.25)	0. (0)	0.26 (0.71)	0.39 (1.13)
Control errors – RVF	0.33 (0.78)	0.13 (0.53)	0.53 (0.94)	0.13 (0.53)	0.53 (0.94)
Control errors – Overall	0.33 (0.56)	0.39 (0.65)	0.26 (0.47)	0.2 (0.42)	0.46 (0.66)

Means and standard deviations for 2D:4D measures, counting and control task reaction times and percentage error scores in females only.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.978 (0.03)	0.953 (0.02)	1.001 (0.03)	0.963 (0.02)	0.992 (0.04)
Left hand 2D:4D	0.983 (0.03)	0.965 (0.02)	0.999 (0.03)	0.959 (0.01)	1.005 (0.03)
Counting RT – LVF	1108.82 (252.92)	1166.05 (296.82)	1055.4 (199.47)	1214.04 (298.8)	1010.62 (153.71)
Counting RT – RVF	1132.76 (277.65)	1204.43 (331.2)	1065.87 (205.89)	1234.71 (337.23)	1037.6 (168.76)
Counting RT – Overall	1110.66 (263.72)	1178.45 (313.47)	1047.4 (197.41)	1214.29 (316.79)	1013.95 (158.58)
Control RT – LVF	556.05 (89.47)	549.96 (80.31)	561.73 (99.75)	545.79 (81.46)	565.63 (98.21)
Control RT – RVF	552.03 (86.38)	546.75 (82.6)	556.97 (92.38)	546.39 (81.74)	557.3 (93.05)
Control RT – Overall	554.5 (87.01)	548.68 (80.63)	559.93 (95.07)	546.54 (80.99)	561.93 (94.49)
Counting errors – LVF	13.61 (13.05)	9.14 (7.16)	17.77 (15.95)	13.35 (14.54)	13.84 (12)
Counting errors – RVF	15.65 (12.51)	12.91 (12.68)	18.21 (12.22)	13.2 (14.55)	17.93 (10.24)
Counting errors – Overall	14.61 (11.76)	10.95 (8.94)	18.03 (13.29)	13.2 (13.37)	15.92 (10.34)
Control errors – LVF	0.51 (0.92)	0.45 (0.89)	0.57 (0.98)	0.6 (0.98)	0.43 (0.89)
Control errors – RVF	0.51 (1.21)	0.6 (1.27)	0.43 (1.19)	0.45 (0.89)	0.56 (1.48)
Control errors – Overall	0.51 (0.87)	0.52 (0.79)	0.5 (0.97)	0.52 (0.8)	0.49 (0.97)

Number comparison

Means and standard deviations for 2D:4D measures, number comparison and control task reaction times and percentage error scores in the entire sample, i.e. including both males and females.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.972 (0.03)	0.947 (0.02)	0.995 (0.02)	0.955 (0.02)	0.988 (0.03)
Left hand 2D:4D	0.976 (0.03)	0.961 (0.02)	0.988 (0.03)	0.955 (0.01)	0.995 (0.02)
No. Comp RT – LVF	978.11 (234.29)	1015.74 (274.76)	944.77 (189.45)	971.55 (222.08)	984.28 (248.41)
No. Comp RT – RVF	985.3 (237.47)	1032.85 (271.96)	943.19 (196.61)	984.36 (207.53)	986.19 (265.73)
No. Comp RT – Overall	982.05 (234.81)	1026.82 (271.88)	942.39 (191.68)	981.5 (215.41)	982.56 (254.98)
Control RT – LVF	787.95 (145.28)	779.9 (150.1)	795.09 (142.69)	770.22 (126.27)	804.65 (161.26)
Control RT – RVF	777.52 (138.52)	769.19 (147.69)	784.89 (131.6)	753.78 (120.14)	799.85 (152.23)
Control RT – Overall	783.11 (143.35)	775.29 (150.6)	790.03 (138.46)	761.8 (125.7)	803.16 (157.42)
No. Comp errors – LVF	20.23 (7.73)	18.88 (7.31)	21.43 (8.01)	19.77 (8.18)	20.67 (7.38)
No. Comp errors – RVF	20.52 (7.68)	19.79 (7.71)	21.17 (7.7)	20.69 (8.56)	20.36 (6.87)
No. Comp errors – Overall	20.23 (7.15)	19.33 (7.04)	21.03 (7.25)	20.23 (8.01)	20.23 (6.36)
Control errors – LVF	3.51 (4.13)	3.36 (3.66)	3.64 (4.55)	3.62 (3.46)	3.41 (4.72)
Control errors – RVF	4.67 (4.77)	4.53 (5.13)	4.8 (4.51)	4.85 (5.17)	4.51 (4.44)
Control errors – Overall	4.09 (3.75)	3.95 (3.74)	4.21 (3.81)	4.24 (3.68)	3.94 (3.87)

Means and standard deviations for 2D:4D measures, number comparison and control task reaction times and percentage error scores in males only.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.965 (0.03)	0.944 (0.01)	0.985 (0.01)	0.952 (0.02)	0.977 (0.02(
Left hand 2D:4D	0.969 (0.2)	0.96 (0.02)	0.978 (0.02)	0.951 (0.01)	0.986 (0.01)
No. Comp RT – LVF	1037.23 (256.93)	1079.91 (300.53)	997.06 (209.33)	1015 (205.65)	1058.15 (302.36)
No. Comp RT – RVF	1057.08 (274.3)	1113.91 (317.52)	1003.59 (222.93)	1036.28 (219.25)	1076.65 (323.38)
No. Comp RT – Overall	1047.3 (263.45)	1097.75 (304.9)	999.82 (216.14)	1029.75 (209.63)	1063.82 (311.46)
Control RT – LVF	828.06 (154.66)	818.59 (168.95)	836.97 (144.59)	811.25 (125.97)	843.88 (180.04)
Control RT – RVF	819.95 (150.63)	814.81 (171.89)	824.79 (132.77)	801.13 (127.03)	837.68 (171.94)
Control RT – Overall	825.02 (154.06)	816.94 (169.47)	832.62 (142.9)	806 (126.09)	842.91 (178.5)
No. Comp errors – LVF	19.01 (8.07)	17.02 (7.29)	20.9 (8.53)	17.97 (8.44)	20 (7.84)
No. Comp errors – RVF	18.98 (7.97)	17.94 (7.04)	19.97 (8.86)	18.51 (8.34)	19.43 (7.84)
No. Comp errors – Overall	18.71 (7.3)	17.48 (6.71)	19.87 (7.84)	18.25 (7.94)	19.14 (6.86)
Control errors – LVF	3.41 (4.23)	3.16 (3.65)	3.64 (4.82)	3.35 (3.44)	3.46 (4.98)
Control errors – RVF	4.39 (4.81)	4.2 (5.57)	4.58 (4.14)	4.23 (5.38)	4.55 (4.38)
Control errors – Overall	3.9 (3.58)	3.7 (3.63)	4.08 (3.64)	3.82 (3.48)	3.97 (3.79)

Means and standard deviations for 2D:4D measures, number comparison and control task reaction times and percentage error scores in females only.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.979 (0.04)	0.95 (0.02)	1.003 (0.03)	0.958 (0.03)	0.999 (0.03)
Left hand 2D:4D	0.982 (0.03)	.963 (0.02)	0.998 (0.03)	0.958 (0.02)	1 (0.03)
No. Comp RT – LVF	918.98 (195.63)	947.3 (235.1)	895.39 (158.83)	928.09 (235.81)	910.41 (155.74)
No. Comp RT – RVF	913.53 (169.34)	946.4 (186.98)	886.14 (153.09)	932.44 (187.57)	895.74 (153.86)
No. Comp RT – Overall	916.79 (183.88)	951.17 (216.59)	888.14 (151.99)	933.25 (216.76)	901.29 (151.81)
Control RT – LVF	747.85 (125.04)	738.63 (119.01)	755.53 (132.79)	729.19 (116.16)	765.41 (133.96)
Control RT – RVF	735.08 (112.12)	720.53 (100.83)	747.19 (122.26)	706.44 (94.4)	762.03 (123.27)
Control RT – Overall	741.2 (119.94)	730.87 (117.18)	749.81 (124.89)	717.59 (112.19)	763.41 (126.05)
No. Comp errors – LVF	21.45 (7.3)	20.87 (7.02)	21.94 (7.69)	21.56 (7.77)	21.35 (7.07)
No. Comp errors – RVF	22.05 (7.16)	21.76 (8.13)	22.3 (6.48)	22.87 (8.47)	21.29 (5.84)
No. Comp errors – Overall	21.75 (6.76)	21.31 (7.06)	22.12 (6.68)	22.21 (7.81)	21.32 (5.82)
Control errors – LVF	3.61 (4.08)	3.57 (3.79)	3.64 (4.42)	3.88 (3.57)	3.35 (4.61)
Control errors – RVF	4.95 (4.79)	4.88 (4.77)	5.01 (4.95)	5.46 (5.06)	4.46 (4.64)
Control errors – Overall	4.28 (3.96)	4.22 (3.97)	4.33 (4.06)	4.67 (3.93)	3.91 (4.06)

SNARC

Means and standard deviations for 2D:4D measures, and average SNARC task and Vowel consonant task regression weights in the entire sample, i.e. including both males and females.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.971 (0.03)	0.948 (0.02)	0.996 (0.02)	0.959 (0.03)	0.983 (0.03)
Left hand 2D:4D	0.976 (0.03)	0.963 (0.02)	0.989 (0.02)	.956 (0.02)	0.995 (0.02)
SNARC	-0.19 (0.35)	-0.2 (0.37)	-0.17 (0.33)	-0.26 (0.37)	-0.12 (0.32)
Vowel-consonant	-.006 (0.32)	0.02 (0.34)	-0.15 (0.29)	-0.041 (0.35)	-0.08 (0.3)

Means and standard deviations for 2D:4D measures, and average SNARC task and Vowel consonant task regression weights in males only.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.962 (0.03)	0.939 (0.02)	0.984 (0.01)	0.95 (0.02)	0.972 (0.03)
Left hand 2D:4D	0.968 (0.02)	0.958 (0.02)	0.978 (0.02)	0.949 (0.01)	0.984 (0.01)
SNARC	-0.19 (0.35)	-0.2 (0.36)	-0.18 (0.36)	-0.28 (0.34)	-0.12 (0.36)
Vowel-consonant	-0.1 (0.29)	-0.01 (0.31)	-0.19 (0.24)	-0.13 (0.34)	-0.07 (0.25)

Means and standard deviations for 2D:4D measures, and average SNARC task and Vowel consonant task regression weights in females only.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.981 (0.04)	0.957 (0.02)	1.011 (0.02)	0.967 (0.03)	0.996 (0.03)
Left hand 2D:4D	0.984 (0.03)	0.969 (0.02)	1.003 (0.03)	0.962 (0.01)	1.008 (0.03)
SNARC	-0.18 (0.35)	-0.21 (0.39)	-0.153 (0.3)	-0.24 (0.4)	-0.11 (0.28)
Vowel-consonant	-0.01 (0.36)	0.04 (0.37)	-0.088 (0.34)	0.04 (0.36)	-0.08 (0.36)

Assessment of normality

Results from the Kolmogorov-Smirnov analysis to explore normality in the subitizing data set.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RHAv2D4D	.063	63	.200*	.980	63	.412
LHAv2D4D	.111	63	.056	.946	63	.009
RT2overall	.112	63	.049	.950	63	.013
RT3overall	.131	63	.009	.921	63	.001
RT4overall	.080	63	.200*	.969	63	.117
SubRTleft	.103	63	.097	.950	63	.012
SubRTright	.106	63	.073	.948	63	.010
SubRToverall	.115	63	.039	.954	63	.019
RTredoverall	.129	63	.010	.949	63	.011
RTbluoverall	.088	63	.200*	.965	63	.071
RTyeloverall	.120	63	.024	.929	63	.001
ConRTleft	.117	63	.032	.966	63	.077
ConRTright	.125	63	.016	.949	63	.011
ConRToverall	.120	63	.025	.956	63	.025
Err2overall	.335	63	.000	.533	63	.000
Err3overall	.225	63	.000	.750	63	.000
Err4overall	.155	63	.001	.841	63	.000
suberrle	.146	63	.002	.889	63	.000
suberrri	.164	63	.000	.814	63	.000
suberov	.157	63	.001	.838	63	.000
Errredoverall	.514	63	.000	.418	63	.000
Errbluoverall	.478	63	.000	.527	63	.000
Erryeloverall	.512	63	.000	.397	63	.000
conerrle	.458	63	.000	.573	63	.000
conerrri	.464	63	.000	.497	63	.000
conerrov	.345	63	.000	.659	63	.000

*. This is a lower bound of the true significance.

Results from the Kolmogorov-Smirnov analysis to explore normality in the counting data set.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RHAv2D4D	.065	61	.200*	.970	61	.144
LHAv2D4D	.111	61	. 061	.946	61	.010
RT6overall	.129	61	.013	.933	61	.002
RT7overall	.103	61	.167	.961	61	.049
RT8overall	.161	61	.000	.904	61	.000
CountRTleft	.097	61	.200*	.966	61	.090
CountRTright	.101	61	.197	.960	61	.044
CountRToverall	.100	61	.200*	.954	61	.023
RTredoverall	.178	61	.000	.837	61	.000
RTbluoverall	.161	61	.000	.825	61	.000
RTyeloverall	.215	61	.000	.854	61	.000
ConRTleft	.179	61	.000	.837	61	.000
ConRTright	.180	61	.000	.875	61	.000
ConRToverall	.192	61	.000	.854	61	.000
Err6overall	.185	61	.000	.839	61	.000
Err7overall	.126	61	.017	.900	61	.000
Err8overall	.114	61	.047	.922	61	.001
CountErrleft	.133	61	.009	.885	61	.000
CountErrright	.131	61	.011	.932	61	.002
CountErroverall	.129	61	.013	.908	61	.000
Errredoverall	.525	61	.000	.372	61	.000
Errbluoverall	.493	61	.000	.414	61	.000
Erryeloverall	.532	61	.000	.268	61	.000
ConErrorsleft	.492	61	.000	.489	61	.000
ConErrorsright	.496	61	.000	.464	61	.000
ConErrorsoverall	.420	61	.000	.621	61	.000

*. This is a lower bound of the true significance.

Results from the Kolmogorov-Smirnov analysis to explore normality in the number comparison data set.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RHAv2D4D	.060	66	.200*	.982	66	.469
LHAv2D4D	.103	66	.080	.951	66	.011
RTnc0.57overall	.133	66	.006	.913	66	.000
RTnc0.67overall	.126	66	.011	.910	66	.000
RTnc0.8overall	.130	66	.008	.916	66	.000
NocompRTleft	.116	66	.029	.910	66	.000
NocompRTright	.112	66	.040	.916	66	.000
NocompRToverall	.124	66	.014	.912	66	.000
RTcon0.57overall	.100	66	.096	.921	66	.000
RTcon0.67overall	.099	66	.181	.919	66	.000
RTcon0.8overall	.112	66	.038	.916	66	.000
ConRTleft	.115	66	.031	.912	66	.000
ConRTright	.143	66	.002	.913	66	.000
ConRToverall	.117	66	.024	.913	66	.000
Errnc0.57overall	.209	66	.000	.812	66	.000
Errnc0.67overall	.097	66	.200*	.946	66	.007
Errnc0.8overall	.079	66	.200*	.983	66	.519
NocompErrorsleft	.083	66	.200*	.980	66	.369
NocompErrorsright	.088	66	.200*	.976	66	.230
NocompErrorsoverall	.083	66	.200*	.966	66	.069
Errcon0.57overall	.466	66	.000	.481	66	.000
Errcon0.67overall	.440	66	.000	.602	66	.000
Errcon0.8overall	.207	66	.000	.883	66	.000
ConErrorsleft	.213	66	.000	.797	66	.000
ConErrorsright	.188	66	.000	.860	66	.000
ConErrorsoverall	.206	66	.000	.874	66	.000

*. This is a lower bound of the true significance.

Results from the Kolmogorov-Smirnov analysis to explore normality in the SNARC data set.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RHAv2D4D	.056	67	.200*	.981	67	.395
LHAv2D4D	.105	67	.065	.952	67	.012
SNARCweight	.063	67	.200*	.986	67	.675
VCweight	.070	67	.200*	.979	67	.321
SNAmeanrt	.073	67	.200*	.963	67	.046
VCmeanrt	.105	67	.064	.949	67	.008
errSNARC	.198	67	.000	.894	67	.000
errVC	.163	67	.000	.845	67	.000

*. This is a lower bound of the true significance.

Section 6.4.1

Non-significant Z and p values from the Wilcoxn signed ranks analysis conducted in order to explore any differences in reaction time and percentage error scores between stimuli on the subitizing control task.

Comparison	**Z**	**p**
Red vs. Yellow reaction times	-0.668	0.504
Red vs. Blue reaction times	-1.034	0.301
Yellow vs. Blue reaction times	-0.342	0.732
Red vs. Yellow percentage error	-1.667	0.096
Red vs. Blue percentage error	-0.243	0.808
Yellow vs. Blue percentage error	-1.376	0.169

Spearman's correlation coefficients (ρ) demonstrating non-significant speed/accuracy associations for performance on the subitizing and control task (n = 63).

Analysis	ρ	**p**
Subitizing left visual field	0.197	0.121
Colour left visual field	0.055	0.669
Colour right visual field	0.157	0.219
Colour overall	0.132	0.302

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (subitizing vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on subitizing reaction times (all df = 1,59).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Visual field	<0.001	0.999	356.818	<0.00001	0.05
Task x Sex interaction	0.001	0.972	5457.395	0.00002	0.05
Task x Visual field interaction	3.652	0.061	422.374	0.058	0.468
Sex x Visual field interaction	0.043	0.837	356.818	0.001	0.055
Task x Sex x Visual field interaction	0.911	0.344	422.374	0.015	0.155

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (subitizing vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on subitizing percentage error scores (all df = 1,59).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Sex	0.457	0.502	58.133	0.008	0.102
Main effect Visual field	0.266	0.608	14.343	0.004	0.08
Task x Sex interaction	0.726	0.398	54.853	0.012	0.134
Task x Visual field interaction	0.365	0.548	12.394	0.006	0.091
Sex x Visual field interaction	0.241	0.625	14.342	0.004	0.077
Task x Sex x Visual field interaction	1.393	0.243	12.394	0.023	0.213

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (subitizing vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on subitizing reaction times (all df = 1,59).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Visual field	0.001	0.971	359.62	0.00003	0.05
Task x Sex interaction	0.001	0.978	5373.689	0.00001	0.05
Task x Visual field interaction	3.911	0.053	403.962	0.062	0.494
Sex x Visual field interaction	0.069	0.793	359.62	0.001	0.058
Task x Sex x Visual field interaction	1.153	0.287	403.962	0.019	0.184

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (subitizing vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on subitizing percentage error scores (all df = 1,59).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Sex	0.538	0.466	54.14	0.009	0.112
Main effect Visual field	0.231	0.633	14.425	0.004	0.076
Task x Sex interaction	0.878	0.353	50.578	0.015	0.152
Task x Visual field interaction	0.289	0.593	12.646	0.005	0.083
Sex x Visual field interaction	0.24	0.626	14.425	0.004	0.077
Task x Sex x Visual field interaction	1.403	0.241	12.646	0.023	0.214

Means and standard deviations of right and left hand 2D:4D and subitizing Z-scores for the reanalysis of the relationship between 2D:4D and subitizing (data combined from experiments 2-4).

	Males N = 97	Females N = 112	Overall N = 209
RH 2D:4D	0.967 (0.03)	0.984 (0.03)	0.976 (0.03)
LH 2D:4D	0.968 (0.03)	0.981 (0.03)	0.975 (0.03)
Subitizing RT Z-scores	0.1 (1.2)	-0.1 (0.8)	0 (1)

Results from the Kolmogorov-Smirnov analysis to explore normality of variable included of the reanalysis of the relationship between 2D:4D and subitizing (data combined from experiments 2-4).

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Final right hand 2D:4D	.055	209	.200*	.994	209	.590
Final left hand 2D:4D	.041	209	.200*	.991	209	.191
Zsubrtoverall	.074	209	.007	.943	209	.000

*. This is a lower bound of the true significance.

Section 6.4.2

Non-significant Z and p values from the Wilcoxn signed ranks analysis conducted in order to explore any differences in reaction time and percentage error scores between stimuli on the counting control task.

Comparison	**Z**	**p**
Red vs. Yellow reaction times	-0.018	0.985
Red vs. Blue reaction times	-1.163	0.245
Yellow vs. Blue reaction times	-1.237	0.216
Red vs. Yellow percentage error	-1.125	0.261
Red vs. Blue percentage error	-0.325	0.745
Yellow vs. Blue percentage error	-1.211	0.226

Spearman's correlation coefficients demonstrating speed/accuracy associations for performance on the counting control task (n = 61).

Analysis	***ρ***	**p**
Colour left visual field	-0.049	0.709
Colour right visual field	0.189	0.144
Colour overall	0.044	0.737

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (coutning vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on counting reaction times (all df = 1,57).

Effect	**F**	**p**	**MSe**	ηp^2	**1-β**
Main effect Sex	0.138	0.712	104092.3	0.002	0.065
Main effect Visual field	1.286	0.261	2428.998	0.022	0.2
Task x Sex interaction	1.111	0.296	77524.75	0.019	0.179
Sex x Visual field interaction	0.232	0.632	2428.998	0.004	0.076
Task x Sex x Visual field interaction	0.172	0.68	2520.95	0.003	0.069

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (coutning vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on counting percentage error scores (all df = 1,57).

Effect	**F**	**p**	**MSe**	ηp^2	**1-β**
Main effect Sex	0.456	0.502	150.818	0.008	0.102
Task x Sex interaction	0.657	0.421	143.491	0.011	0.125
Sex x Visual field interaction	0.053	0.819	19.705	0.001	0.056
Task x Sex x Visual field interaction	0.05	0.824	20.634	0.001	0.056

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (counting vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on counting reaction times (all df = 1,57).

Effect	**F**	**p**	**MSe**	ηp^2	**1-β**
Main effect Sex	0.154	0.697	101114.8	0.003	0.067
Main effect Visual field	1.258	0.267	2403.775	0.022	0.197
Task x Sex interaction	1.224	0.273	73030.731	0.021	0.193
Task x Visual field	3.278	0.076	2394.009	0.054	0.429
Sex x Visual field interaction	0.217	0.643	2403.775	0.004	0.074
Task x Sex x Visual field interaction	0.157	0.693	2394.069	0.003	0.068

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (counting vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on percentage error scores (all df = 1,57).

Effect	**F**	**p**	**MSe**	ηp^2	**1-β**
Main effect Sex	0.402	0.528	158.967	0.007	0.096
Task x Sex interaction	0.585	0.448	151.627	0.01	0.117
Sex x Visual field interaction	0.087	0.769	19.324	0.002	0.06
Task x Sex x Visual field interaction	0.077	0.782	20.134	0.001	0.059

Results from the Kolmogorov-Smirnov analysis to explore normality of variables included of the reanalysis of the relationship between 2D:4D and counting (data combined from experiments 2 and 4).

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Final right hand 2D:4D	.057	137	.200*	.986	137	.178
Final left hand 2D:4D	.072	137	.079	.975	137	.014
ZCountrtoverall	.056	137	.200*	.991	137	.484
Zcountrtleft	.040	137	.200*	.993	137	.714
Zcountrtright	.049	137	.200*	.992	137	.635

*. This is a lower bound of the true significance.

Means and standard deviations of right and left hand 2D:4D and subitizing Z-scores (for stimuli presented to both the right visual field (RFV), left visual field (LVF) and overall) for the reanalysis of the relationship between 2D:4D and subitizing (data combined from experiments 2-4).

	Males N = 62	Females N = 75	Overall N = 137
RH 2D:4D	0.97 (0.03)	0.986 (0.03)	0.979 (0.03)
LH 2D:4D	0.974 (0.03)	0.987 (0.03)	0.981 (0.03)
Subitizing RT Z-scores – Overall	-0.1 (1.1)	0.1 (0.9)	0 (1)
Subitizing RT Z-scores – LVF	-0.1 (1.1)	0.1 (0.9)	0 (1)
Subitizing RT Z-scores - RVF	-0.2 (1.1)	0.1 (0.9)	0 (1)

Section 6.4.3

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (number comparison vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on number comparison reaction times (all df = 1,62).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Visual field	0.102	0.75	1445.129	0.002	0.061
Task x Sex interaction	1.721	0.194	20618.702	0.027	0.253
Sex x Visual field interaction	0.987	0.342	1445.129	0.041	0.361
Task x Sex x Visual field interaction	1.805	0.184	918.996	0.028	0.263

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (number comparison vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on number comparison percentage error scores (all df = 1,62).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Visual field	2.835	0.097	12.601	0.044	0.381
Task x Sex interaction	2.341	0.131	39.836	0.036	0.325
Sex x Visual field interaction	0.313	0.578	12.601	0.005	0.085
Task x Sex x Visual field interaction	0.016	0.899	19.398	0.0002	0.052

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (number comparison vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on number comparison reaction times (all df = 1,62).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Visual field	0.117	0.733	1487.901	0.002	0.063
Task x Sex interaction	1.64	0.205	223322.521	0.026	0.243
Sex x Visual field interaction	2.486	0.12	1487.901	0.039	0.342
Task x Sex x Visual field interaction	1.878	0.175	891.449	0.029	0.271

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (number comparison vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on number comparison percentage error scores (all df = 1,62).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Sex	1.887	0.174	88.333	0.03	0.272
Main effect Visual field	2.827	0.098	12.598	0.044	0.38
Task x Sex interaction	2.338	0.131	40.536	0.036	0.325
Task x Visual field interaction	0.627	0.431	19.425	0.01	0.122
Sex x Visual field interaction	0.329	0.568	12.598	0.005	0.087
Task x Sex x Visual field interaction	0.016	0.901	19.425	0.0003	0.052

Section 6.4.3

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 3-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, 2D:4D (low vs. high), task (SNARC vs. control) and sex (males vs. females) on calculated regression weights (all df = 1,63).

Analysis	**Effect**	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Right Hand 2D:4D	Main effect Sex	0.506	0.48	0.124	0.008	0.108
	Task x Sex interaction	0.315	0.576	0.105	0.005	0.086
Left Hand 2D:4D	Main effect Sex	0.338	0.563	0.123	0.011	0.134
	Task x Sex interaction	0.725	0.398	0.105	0.005	0.088

Appendix 5

Additional statistical information relating Chapter 7

Section 7.2.3

Average 2D:4D values and t-test analysis of group differences in 2D:4D for high and low 2D:4D groups for subitizing and number comparison task data sets.

		Mean RH 2D:4D		Group comparisons		Mean LH 2D:4D		Group comparisons	
		Low	High	t	p	Low	High	t	p
Subitizing	Males (n = 26)	0.93	0.97	5.334	<0.001	0.92	0.96	6.01	<0.001
	Females (n = 28)	0.94	0.99	7.411	<0.001	0.93	0.98	6.938	<0.001
Number Comparison	Males (n = 21)	0.93	0.97	4.296	<0.001	0.92	0.97	4.846	<0.001
	Females (n = 29)	0.93	0.99	8.5	<0.001	0.93	0.97	6.111	<0.001

Section 7.3

Subitizing

Means and standard deviations for 2D:4D measures, subitizing and control task reaction times and percentage error scores in the entire sample, i.e. including both males and females.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.956 (0.03)	0.933 (0.018)	0.98 (0.02)	0.945 (0.03)	0.968 (0.03)
Left hand 2D:4D	0.948 (0.03)	0.936 (0.024)	0.96 (0.03)	0.925 (0.01)	0.971 (0.02)
Subitizing RT – LVF	752.65 (189.01)	772.99 (203.36)	732.31 (174.96)	768.19 (181.25)	737.12 (198.68)
Subitizing RT – RVF	767.36 (198.25)	744.7 (182.04)	790.02 (214.27)	756.28 (158.01)	778.45 (234.3)
Subitizing RT – Overall	747.04 (179.22)	748.38 (183.55)	745.7 (178.28)	756.77 (162.81)	737.31 (196.91)
Control RT – LVF	588.44 (134.84)	581.82 (102.59)	595.07 (162.62)	604.31 (105.55)	572.57 (159.37)
Control RT – RVF	577.46 (143.35)	565.8 (90.76)	589.12 (182.68)	611.27 (133.5)	543.66 (147.29)
Control RT – Overall	575.32 (131.24)	565.56 (96.37)	585.09 (160.08)	597.25 (109.73)	553.4 (148.57)
Subitizing errors – LVF	15.29 (11.1)	16.45 (11.78)	14.13 (10.47)	16.55 (11.12)	14.03 (11.14)
Subitizing errors – RVF	15.24 (10.29)	16.12 (11.11)	14.37 (9.54)	16.21 (10.15)	14.27 (10.54)
Subitizing errors – Overall	15.29 (11.1)	16.45 (11.78)	14.13 (10.47)	16.55 (11.12)	14.03 (11.14)
Control errors – LVF	4.38 (7.65)	4.26 (7.39)	4.5 (8.04)	4.25 (7.4)	4.5 (8.03)
Control errors – RVF	4.09 (6.5)	4.46 (7.04)	3.71 (6.03)	4.35 (6.88)	3.82 (6.22)
Control errors – Overall	4.29 (6.6)	4.47 (6.93)	4.11 (6.39)	4.4 (6.87)	4.18 (6.45)

Means and standard deviations for 2D:4D measures, subitizing and control task reaction times and percentage error scores in males only.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.948 (0.03)	0.928 (0.02)	0.968 (0.02)	0.9414 (0.03)	0.954 (0.03)
Left hand 2D:4D	0.94 (0.03)	0.93 (0.02)	0.949 (0.04)	0.917 (0.01)	0.962 (0.03)
Subitizing RT – LVF	706.19 (143.69)	741.92 (163.18)	670.46 (116.71)	765.31 (158.11)	647.08 (102.19)
Subitizing RT – RVF	716.51 (160.86)	710.86 (112.68)	722.17 (202.84)	714.81 (77.84)	718.22 (218.73)
Subitizing RT – Overall	695.65 (123.39)	712.13 (109.91)	679.18 (138.03)	730.92 (103.27)	660.38 (135.5)
Control RT – LVF	578.33 (151.5)	571.44 (90.89)	585.23 (198.62)	595.27 (112.07)	561.4 (186.11)
Control RT – RVF	564.92 (162.53)	565.15 (91.82)	564.68 (215.87)	607.05 (137.84)	522.78 (179.4)
Control RT – Overall	566.1 (152.82)	562.76 (96.88)	569.44 (198.1)	594.28 (124.81)	537.91 (177.08)
Subitizing errors – LVF	15.83 (12.5)	13.77 (13.52)	17.89 (11.55)	16.06 (12.54)	15.6 (12.97)
Subitizing errors – RVF	15.28 (11.82)	13.3 (12.84)	17.27 (10.83)	13.92 (11.16)	16.64 (12.74)
Subitizing errors – Overall	15.83 (12.5)	13.77 (13.52)	17.89 (11.55)	16.06 (12.54)	15.6 (12.97)
Control errors – LVF	6.43 (9.37)	6.5 (9.53)	6.37 (9.6)	6.27 (9.53)	6.59 (9.6)
Control errors – RVF	5.4 (8.48)	4.9 (9.49)	5.91 (7.7)	5.52 (9.37)	5.29 (7.87)
Control errors – Overall	6 (8.54)	5.86 (9.47)	6.15 (7.9)	6.01 (9.36)	6 (8.03)

Means and standard deviations for 2D:4D measures, subitizing and control task reaction times and percentage error scores in females only.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.965 (0.03)	0.938 (0.02)	0.991 (0.02)	0.949 (0.03)	0.981 (0.02)
Left hand 2D:4D	0.955 (0.03)	0.941 (0.03)	0.969 (0.03)	0.931 (0.01)	0.979 (0.02)
Subitizing RT – LVF	795.8 (216.79)	801.85 (237.2)	789.75 (203.15)	770.86 (206.41)	820.74 (231.66)
Subitizing RT – RVF	814.58 (219.98)	776.13 (228.8)	853.04 (212.06)	794.79 (202.54)	834.38 (242.15)
Subitizing RT – Overall	794.76 (209.89)	782.04 (231.72)	807.48 (193.52)	780.76 (204.65)	808.75 (221.15)
Control RT – LVF	597.83 (119.36)	591.46 (114.96)	604.2 (127.61)	612.71 (102.61)	582.95 (136.32)
Control RT – RVF	589.11 (124.85)	566.4 (93.22)	611.82 (150.25)	615.18 (134.44)	563.05 (113.31)
Control RT – Overall	583.89 (109.67)	568.15 (99.47)	599.62 (120.64)	600 (98.4)	567.77 (121.41)
Subitizing errors – LVF	14.8 (9.83)	18.95 (9.73)	10.65 (8.29)	17.01 (10.09)	12.58 (9.38)
Subitizing errors – RVF	15.21 (8.87)	18.75 (8.89)	11.67 (7.57)	18.34 (8.98)	12.07 (7.83)
Subitizing errors – Overall	14.8 (9.83)	18.95 (9.73)	10.65 (8.29)	17.01 (10.09)	12.58 (9.38)
Control errors – LVF	2.47 (5.08)	2.18 (3.99)	2.76 (6.12)	2.38 (4.23)	2.57 (5.97)
Control errors – RVF	2.86 (3.63)	4.05 (3.96)	1.67 (2.93)	3.26 (3.31)	2.47 (4)
Control errors – Overall	2.69 (3.54)	3.17 (3.08)	2.22 (4.01)	2.9 (2.92)	2.49 (4.17)

Number comparison

Means and standard deviations for 2D:4D measures, number comparison and control task reaction times and percentage error scores in the entire sample, i.e. including both males and females.

	Overall	**Low 2D:4D (Right hand split)**	**High 2D:4D (Right hand split)**	**Low 2D:4D (Left hand split)**	**High 2D:4D (Left hand split)**
Right hand 2D:4D	0.953 (0.03)	0.93 (0.01)	0.975 (0.02)	0.937 (0.02)	0.968 (0.03)
Left hand 2D:4D	0.947 (0.03)	0.93 (0.02)	0.964 (0.03)	0.926 (0.01)	0.968 (0.02)
No. Comp RT – LVF	1304.1 (271.11)	1349.24 (305.85)	1268.39 (240.73)	1325.52 (341.44)	1283.61 (186.61)
No. Comp RT – RVF	1267.7 (297.48)	1275.05 (351.42)	1262.72 (254.09)	1276.14 (355.08)	1259.63 (237.69)
No. Comp RT – Overall	1287.49 (273.98)	1308.74 (317.08)	1270.67 (240.51)	1302.05 (339.88)	1273.57 (198.61)
Control RT – LVF	1102.07 (185.15)	1094.93 (193.81)	1108.72 (185.26)	1101.32 (184.97)	1102.78 (189.48)
Control RT – RVF	1071.01 (179.87)	1045.55 (177.39)	1089.26 (185.83)	1063.25 (177.69)	1078.43 (185.61)
Control RT – Overall	1087.02 (180.06)	1070.5 (181.04)	1099.91 (185.73)	1081.95 (175.67)	1091.87 (187.97)
No. Comp errors – LVF	32.83 (14.35)	33.75 (14.53)	32.15 (14.77)	34.51 (16.35)	31.22 (12.3)
No. Comp errors – RVF	34.61 (14.53)	34.08 (17.81)	34.9 (11.52)	34.43 (16.04)	34.79 (13.29)
No. Comp errors – Overall	33.73 (13.2)	33.98 (15.35)	33.49 (11.59)	34.54 (15.02)	32.96 (11.48)
Control errors – LVF	6.56 (7.11)	7.26 (7.32)	6.02 (7.18)	6.5 (6.88)	6.62 (7.48)
Control errors – RVF	8.58 (13.45)	11.93 (18.13)	5.89 (6.45)	11.76 (17.6)	5.54 (6.82)
Control errors – Overall	7.46 (9.17)	9.39 (12.24)	5.93 (4.96)	8.84 (11.8)	6.15 (5.62)

Means and standard deviations for 2D:4D measures, number comparison and control task reaction times and percentage error scores in males only.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.949 (0.03)	0.931 (0.02)	0.965 (0.02)	0.936 (0.02)	0.96 (0.03)
Left hand 2D:4D	0.944 (0.03)	0.925 (0.01)	0.961 (0.03)	0.92 (0.01)	0.966 (0.03)
No. Comp RT – LVF	1270.42 (317.34)	1297.94 (357.01)	1245.65 (294.4)	1234.78 (419.28)	1302.5 (206.7)
No. Comp RT – RVF	1268.34 (368.13)	1269.22 (443.82)	1267.55 (309.76)	1212.89 (498.06)	1318.25 (211.4)
No. Comp RT – Overall	1276.18 (335.78)	1294 (395.89)	1260.15 (292.54)	1225.94 (455.47)	1321.4 (190.5)
Control RT – LVF	1044.26 (196.37)	995.17 (246.95)	1088.45 (135.4)	1004.67 (248.7)	1079.9 (138.44)
Control RT – RVF	1009.53 (187.5)	949.33 (205.85)	1063.7 (160.51)	988.72 (231.37)	1028.25 (148.02)
Control RT – Overall	1030.37 (193.3)	975.33 (228.77)	1079.9 (149.86)	997.33 (239.16)	1060.1 (147.71)
No. Comp errors – LVF	32.22 (15.06)	30.27 (13.04)	33.96 (17.18)	33.46 (18.93)	31.09 (11.49)
No. Comp errors – RVF	34.29 (15.58)	32.58 (18.86)	35.83 (12.81)	36.2 (17.05)	32.57 (14.84)
No. Comp errors – Overall	33.24 (13.78)	31.6 (15.5)	34.72 (12.7)	34.94 (16.77)	31.72 (11.16)
Control errors – LVF	8.83 (8.97)	8.49 (9.82)	9.14 (8.66)	8.06 (9.61)	9.52 (8.81)
Control errors – RVF	11.91 (17.8)	17.36 (24.04)	7 (7.96)	18.31 (23.6)	6.14 (7.77)
Control errors – Overall	10.43 (12.31)	13.2 (17.1)	7.94 (5.36)	13.28 (16.67)	7.86 (6.37)

Means and standard deviations for 2D:4D measures, number comparison and control task reaction times and percentage error scores in females only.

	Overall	Low 2D:4D (Right hand split)	High 2D:4D (Right hand split)	Low 2D:4D (Left hand split)	High 2D:4D (Left hand split)
Right hand 2D:4D	0.957 (0.03)	0.929 (0.01)	0.983 (0.02)	0.938 (0.03)	0.974 (0.03)
Left hand 2D:4D	0.95 (0.03)	0.934 (0.02)	0.966 (0.02)	0.93 (0.01)	0.969 (0.02)
No. Comp RT – LVF	1328.71 (235.29)	1387.71 (271.33)	1285.88 (201.2)	1388.35 (276.34)	1269.08 (176.84)
No. Comp RT – RVF	1267.23 (241.19)	1279.42 (285.03)	1259 (215.34)	1319.92 (224.02)	1214.54 (254.93)
No. Comp RT – Overall	1295.75 (225.32)	1319.79 (261.74)	1278.77 (204.21)	1354.73 (237.71)	1236.77 (204.27)
Control RT – LVF	1144.31 (167.86)	1169.75 (98.44)	1124.31 (220.38)	1168.23 (82.38)	1120.38 (225.11)
Control RT – RVF	1115.94 (163.27)	1117.71 (114.64)	1108.92 (207.42)	1114.85 (111.88)	1117.04 (207.41)
Control RT – Overall	1128.42 (161.04)	1141.88 (92.52)	1115.31 (214.03)	1140.54 (82.39)	1116.31 (216.62)
No. Comp errors – LVF	33.28 (14.1)	36.36 (15.58)	30.76 (13.18)	35.24 (15.08)	31.33 (13.36)
No. Comp errors – RVF	34.85 (14.02)	35.21 (17.74)	34.2 (10.91)	33.21 (15.89)	36.5 (12.3)
No. Comp errors – Overall	34.09 (13.02)	35.76 (15.66)	32.54 (11.1)	34.26 (14.4)	33.91 (12.08)
Control errors – LVF	4.9 (4.94)	6.35 (5.01)	3.61 (4.89)	5.41 (4.26)	4.4 (5.67)
Control errors – RVF	6.15 (8.71)	7.86 (11.61)	5.04 (5.19)	7.22 (10.77)	5.08 (6.28)
Control errors – Overall	5.29 (5.22)	6.54 (6.26)	4.38 (4.21)	5.76 (5.76)	4.83 (4.82)

Assessment of normality

Results from the Kolmogorov-Smirnov analysis to explore normality in the subitizing data set.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RH2D4D	.068	54	.200*	.978	54	.422
LH2D4D	.094	54	.200*	.943	54	.012
RT2overall	.097	54	.200*	.930	54	.004
RT3overall	.146	54	.006	.936	54	.006
RT4overall	.196	54	.000	.799	54	.000
SubleftRT	.154	54	.003	.889	54	.000
SubrightRT	.125	54	.036	.942	54	.011
SuboverallRT	.138	54	.012	.925	54	.002
RTbluoverall	.103	54	.200*	.942	54	.011
RTRedoverall	.152	54	.003	.944	54	.014
RTyeloverall	.161	54	.001	.926	54	.003
ConleftRT	.083	54	.200*	.958	54	.055
ConrightRT	.114	54	.078	.911	54	.001
ConoverallRT	.111	54	.096	.958	54	.055
Err2overall	.347	54	.000	.721	54	.000
Err3overall	.155	54	.002	.836	54	.000
Err4overall	.126	54	.032	.920	54	.002
Errsubleft	.115	54	.072	.927	54	.003
Errsubright	.124	54	.039	.951	54	.027
Errsuboverall	.115	54	.072	.927	54	.003
Errbluoverall	.345	54	.000	.640	54	.000
ErrRedoverall	.353	54	.000	.651	54	.000
Erryeloverall	.371	54	.000	.552	54	.000
Errconleft	.309	54	.000	.632	54	.000
Errconright	.265	54	.000	.639	54	.000
Errconoverall	.258	54	.000	.661	54	.000

*. This is a lower bound of the true significance.

Results from the Kolmogorov-Smirnov analysis to explore normality in the subitizing data set.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RH2D4D	.095	44	.200*	.958	44	.113
LH2D4D	.123	44	.090	.946	44	.038
RTnc0.56overall	.089	44	.200*	.980	44	.641
RTnc0.67overall	.070	44	.200*	.975	44	.457
RTnc0.8overall	.106	44	.200*	.967	44	.246
NCRTleft	.100	44	.200*	.971	44	.317
NCRTright	.080	44	.200*	.973	44	.383
NCRToverall	.078	44	.200*	.975	44	.455
RTcon0.56overall	.081	44	.200*	.980	44	.623
RTcon0.67overall	.108	44	.200*	.964	44	.190
RTcon0.8overall	.122	44	.096	.973	44	.397
ConRTleft	.120	44	.116	.938	44	.020
ConRTright	.083	44	.200*	.951	44	.062
ConRToverall	.096	44	.200*	.949	44	.049
Errnc0.56overall	.160	44	.006	.935	44	.016
Errnc0.67overall	.101	44	.200*	.968	44	.255
Errnc0.8overall	.088	44	.200*	.990	44	.959
NCerrleft	.116	44	.164	.953	44	.072
NCerrri	.094	44	.200*	.965	44	.207
NCerrov	.111	44	.200*	.958	44	.112
Errcon0.56overall	.298	44	.000	.587	44	.000
Errcon0.67overall	.256	44	.000	.671	44	.000
Errcon0.8overall	.227	44	.000	.798	44	.000
Conerrle	.195	44	.000	.833	44	.000
Conerrri	.259	44	.000	.676	44	.000
Conerrov	.206	44	.000	.737	44	.000

*. This is a lower bound of the true significance.

Section 7.3.1

Non-significant Zand p values from the Wilcoxn signed ranks analysis conducted in order to explore any differences in reaction time and percentage error scores between stimuli on the subitizing control task.

Comparison	**Z**	**p**
Red vs. Yellow reaction times	-0.405	0.686
Red vs. Blue reaction times	-1.184	0.236
Yellow vs. Blue reaction times	-1.158	0.247
Red vs. Yellow percentage error	-0.602	0.547
Red vs. Blue percentage error	-0.377	0.706
Yellow vs. Blue percentage error	-0.411	0.681

Non-significant Spearman's correlation coefficients demonstrating speed/accuracy associations for performance on the subitizing and control task (n = 54).

Analysis	***ρ***	**p**
Colour left visual field	0.244	0.075
Colour right visual field	0.246	0.073
Colour overall	0.238	0.083
Subitizing left visual field	0.048	0.73
Subitizing right visual field	0.093	0.502
Subitizing overall	0.029	0.832

Non-significant Spearman's correlation coefficients demonstrating associations between percentage error scores and age for performance on the subitizing and control task (n = 54).

Analysis	***ρ***	**p**
Subitizing	-0.144	0.298
Control	-0.171	0.216

Results of the post hoc analysis (bonferroni corrected t-tests, $\alpha = 0.0125$) to further investigate the interaction between right hand 2D:4D group and visual field of stimulus presentation on subitizing and subitizing control reaction times.

Comparison	**t**	**df**	**p**
2D:4D group differences in left visual field responses	0.363	52	0.718
2D:4D group differences in right visual field responses	0.86	52	0.394
Visual field differences in low 2D:4D participants	2.076	26	0.048
Visual field differences in high 2D:4D participants	1.786	26	0.086

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (subitizing vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on subitizing reaction times (all df = 1,50).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Sex	2.369	0.13	76150.6	0.045	0.327
Main effect Visual field	0.036	0.849	4490.12	0.001	0.054
Task x Sex interaction	2.733	0.105	25568.3	0.052	0.386
Task x Visual field interaction	1.181	0.282	7491.6	0.023	0.187
Sex x Visual field interaction	0.13	0.72	4490.12	0.003	0.064
Task x Sex x Visual field interaction	0.006	0.937	7491.6	0.0001	0.051

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (subitizing vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on subitizing percentage error scores (all df = 1,50).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Sex	1.003	0.321	194.759	0.02	0.166
Main effect Visual field	0.127	0.723	15.828	0.003	0.064
Task x Sex interaction	0.989	0.325	99.233	0.019	0.164
Task x Visual field interaction	0.073	0.788	11.738	0.001	0.058
Sex x Visual field interaction	1.199	0.279	15.828	0.023	0.189
Task x Sex x Visual field interaction	0.062	0.805	11.738	0.001	0.057

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (subitizing vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on subitizing reaction times (all df = 1,50).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Sex	2.406	0.127	74985.25	0.046	0.331
Main effect Visual field	0.033	0.857	4958.545	0.001	0.054
Task x Sex interaction	2.827	0.099	24717.09	0.054	0.378
Task x Visual field interaction	1.275	0.264	6941.642	0.025	0.198
Sex x Visual field interaction	0.118	0.733	4958.545	0.002	0.063
Task x Sex x Visual field interaction	0.007	0.934	6941.642	0.0001	0.051

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (subitizing vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on subitizing percentage error scores (all df = 1,50).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Sex	0.962	0.331	203.049	0.019	0.161
Main effect Visual field	0.13	0.72	15.463	0.003	0.064
Task x Sex interaction	0.947	0.335	103.665	0.019	0.159
Task x Visual field interaction	0.072	0.79	11.926	0.001	0.058
Sex x Visual field interaction	1.227	0.273	15.463	0.024	0.192
Task x Sex x Visual field interaction	0.061	0.807	11.926	0.001	0.057

Section 7.3.2

Non-significant Spearman's correlation coefficients demonstrating speed/accuracy associations for performance on the number comparison control task (n = 54).

Analysis	***ρ***	**p**
Control left visual field	-0.047	0.758
Control right visual field	-0.266	0.077
Control overall	-0.135	0.378

Pearson's (r) and Spearman's (*ρ*) correlation coefficients demonstrating associations between reaction time and percentage error scores and age (in years) for performance on the number comparison and control task (n = 45).

Analysis	**Statistic**	**p**
Number comparison reaction times	r = 0.155	0.305
Control reaction times	r = -0.186	0.221
Number comparison percentage error scores	r = -0.205	0.177
Control percentage error scores	*ρ* = -0.068	0.136

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power ($1-\beta$) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (number comparison vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on number comparison reaction times (all df = 1,40).

Effect	**F**	**p**	**MSe**	**ηp^2**	**$1-\beta$**
Main effect Sex	1.211	0.278	171670.7	0.029	0.189
Task x Sex interaction	1.116	0.297	51570.7	0.027	0.178
Task x Visual field interaction	0.002	0.969	7086.767	0.00004	0.05
Sex x Visual field interaction	1.221	0.276	8640.292	0.03	0.19
Task x Sex x Visual field interaction	1.643	0.207	7086.767	0.039	0.24

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power ($1-\beta$) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, right hand 2D:4D (low vs. high), task (number comparison vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on number comparison percentage error scores (all df = 1,40).

Effect	**F**	**p**	**MSe**	**ηp^2**	**$1-\beta$**
Main effect Sex	0.442	0.51	353.889	0.011	0.099
Main effect Visual field	2.618	0.114	66.873	0.061	0.352
Task x Sex interaction	1.939	0.171	183.715	0.046	0.274
Task x Visual field interaction	0.1	0.754	70.294	0.002	0.061
Sex x Visual field interaction	0.324	0.572	66.873	0.008	0.086
Task x Sex x Visual field interaction	0.035	0.854	70.294	0.001	0.054

Low and high left hand 2D:4D male and female reaction times on the number comparison and control tasks (averaged across both tasks) for information presented to both the left and right visual fields).

	Males (n = 19)		**Females (n = 26)**	
	Low 2D:4D	**High 2D:4D**	**Low 2D:4D**	**High 2D:4D**
Left	1119.72 (303.7)	1191.2 (133.47)	1278.29 (163.09)	1194.73 (190.49)
Right	1100.81 (337.84)	1173.25 (159.06)	1217.38 (135.39)	1165.79 (212.79)

Subsequent analysis of the 3-way interaction between left hand 2D:4D, sex and visual field on number comparison and control task (averaged across both tasks) percentage error scores. Two separate 2-way ANOVAs were conducted to investigate any possible main effect and/or interaction effects of 2D:4D and visual field on percentage error scores in males and females considered independently.

Analysis	**Effect**	**F value**	**df**	**p**	**MSe**	**ηp^2**	**1-β**
Males	Main effect of 2D:4D	0.599	1,17	0.45	276.139	0.034	0.113
	Main effect of visual field	1.532	1,17	0.233	47.472	0.083	0.215
	2D:4D x Visual field	2.764	1,17	0.115	47.472	0.14	0.348
Females	Main effect of 2D:4D	0.12	1,24	0.732	97.159	0.005	0.063
	Main effect of visual field	1.331	1,24	0.26	19.33	0.053	0.198
	2D:4D x Visual field	1.55	1,24	0.225	19.33	0.061	0.223

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (number comparison vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on number comparison reaction times (all df = 1,41).

Effect	**F**	**p**	**MSe**	**ηp^2**	**1-β**
Main effect Sex	1.227	0.274	164235.7	0.029	0.191
Task x Sex interaction	1.164	0.287	51391.26	0.028	0.184
Task x Visual field interaction	0.002	0.962	6808.208	0.00006	0.05
Sex x Visual field interaction	0.829	0.386	9274.136	0.02	0.144
Task x Sex x Visual field interaction	1.64	0.207	6808.208	0.038	0.24

Non-reported F, p, MSe, effects size (partial eta squared - ηp^2) and power (1-β) values from the 4-way ANOVA analysis conducted in order to explore any main and/or interaction effects of the factors, left hand 2D:4D (low vs. high), task (number comparison vs. control), sex (males vs. females) and visual field of stimulus presentation (left vs. right) on number comparison percentage error scores (all df = 1,41).

Effect	**F**	**p**	**MSe**	ηp^2	**1-β**
Main effect Sex	0.577	0.452	342.74	0.014	0.115
Main effect Visual field	3.085	0.086	61.997	0.07	0.403
Task x Sex interaction	1.922	0.173	186.679	0.045	0.273
Task x Visual field interaction	0.042	0.839	66.697	0.001	0.055
Sex x Visual field interaction	0.329	0.57	61.997	0.008	0.087
Task x Sex x Visual field interaction	0.112	0.74	66.997	0.003	0.062

Appendix 6

Additional statistical information relating Chapter 8

Section 8.2.3

Results from the Kolmogorov-Smirnov analysis to explore normality.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Right hand 2D:4D	.099	56	.200*	.973	56	.242
Left hand 2D:4D	.071	56	.200*	.984	56	.679
Median RT - Simple RT	.216	56	.000	.739	56	.000
Standard age score - Simple RT	.093	56	.200*	.981	56	.512
Stanine - Simple RT	.140	56	.008	.955	56	.036
Pecentage correct - Dot enumeration	.183	56	.000	.838	56	.000
Median RT - Dot enumeration	.057	56	.200*	.980	56	.465
Efficiency measure - Dot enumeation	.080	56	.200*	.984	56	.660
Standard age score - Dot enumeration	.131	56	.018	.964	56	.093
Stanine - Dot enumeration	.157	56	.001	.943	56	.011
Percantage correct - Numerical stroop	.229	56	.000	.891	56	.000
Median RT - Numerical stroop	.155	56	.002	.801	56	.000
Efficiency measure - Numerical stroop	.139	56	.009	.868	56	.000
Standard age score - Numerical stroop	.113	56	.073	.951	56	.023
Stanine - Numerical stroop	.161	56	.001	.940	56	.008
Percentage correct - Addition	.132	56	.016	.914	56	.001
Median RT - Addition	.118	56	.049	.853	56	.000
Efficiency measure - Addition	.078	56	.200*	.923	56	.002
Standard age score - Addition	.175	56	.000	.942	56	.009
Stanine - Addition	.179	56	.000	.899	56	.000

*. This is a lower bound of the true significance.

Means and standard deviations (SD) of male and female scores on all performance measures on the dyscalculia screener for all sub-tests.

	Simple RT		Dot Enumeration		Number Comparison		Arithmetic	
	Males	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
Median RT	418.06 (121.58)	444.94 (169.04)	2799.38 (1095.43)	3268.08 (1235.06)	1256.1 (362.31)	1456.33 (574.09)	5051.22 (3689.91)	5024.94 (2642.25)
Standard Age	94.26 (12.04)	92.63 (12.32)	107.41 (15.19)	101.29 (15.7)	108.21 (15.58)	104.83 (17.4)	97.44 (14.67)	96.21 (14.35)
Stanine	4.12 (1.63)	4 (1.74)	5.53 (1.91)	4.83 (1.69)	5.88 (1.87)	4.96 (1.99)	3.88 (1.72)	3.58 (1.77)
Percentage correct			85.12 (9.46)	85.78 (12.54)	91.27 (14.82)	93.47 (18.35)	77.73 (16.07)	76.34 (16.98)
Efficiency measure			7245.11 (1160.31	6760.94 (1310.58)	4048.13 (425.54)	3902.76 (638.23)	14385.37 (3818.97)	14051.28 (3123.86)

U, z and p values relating to Mann-Whitney *U* analysis of sex differences in performance on the dyscalculia screener (n = 58).

	Simple RT	**Dot Enumeration**	**Number Comparison**	**Arithmetic**
Median RT	U = 346.5 Z = -0.971 p = 0.332	U = 314 Z = -1.484 p = 0.138	U = 297 Z = -1.752 p = 0.08	U = 380.5 Z = 975.5 p = -0.434
Standard Age	U = 365.5 Z = -0.672 p = 0.502	U = 297.5 Z = -1.746 p = 0.081	U = 333.5 Z = -1.177 p = 0.239	U = 352.5 Z = -0.877 p = 0.38
Stanine	U = 392.5 Z = 692.5 p = 0.803	U = 295.5 Z = -1.812 p = 0.07	U = 290.5 Z = -1.885 p = 0.059	U = 357 Z = -0.827 p = 0.408
Percentage correct		U = 367 Z = -0.649 p = 0.516	U = 374.5 Z = -0.532 p = 0.594	U = 391.5 Z = -0.262 p = 0.794
Efficiency measure		U = 307 Z = -1.595 p = 0.111	U = 340 Z = -1.074 p = 0.283	U = 365 Z = -0.821 p = 0.412

Section 8.3.2

Spearman's correlation coefficients (ρ), p values and 1-β values for analysis of the relationship between right hand 2D:4D and standard age, stanine and efficiency measure scores on all subtasks of the Dyscalculia Screener in males (n =34), females (n =24) and the entire data set (n = 58), significant effect indicated in bold.

		Simple RT	Dot enumeration	Number comparison	Addition
Standard age	Males	ρ = -0.242 p = 0.167 1-β = 0.261	ρ = -0.037 p = 0.834 1-β = 0.054	ρ = 0.114 p = 0.522 1-β = 0.093	ρ = 0.104 p = 0.56 1-β = 0.086
	Females	ρ = 0.319 p = 0.138 1-β = 0.297	ρ = -0.37 p = 0.082 1-β = 0.39	ρ = 0.147 p = 0.502 1-β = 0.097	ρ = -0.04 p = 0.857 1-β = 0.053
	Overall	ρ = -0.025 p = 0.856 1-β = 0.054	ρ = -0.202 p = 0.131 1-β = 0.303	ρ = 0.1 p = 0.459 1-β = 0.108	ρ = -0.042 p = 0.757 1-β = 0.06
Stanine	Males	ρ = -0.205 p = 0.245 1-β = 0.199	ρ = -0.122 p = 0.492 1-β = 0.1	ρ = 0.18 p = 0.31 1-β = 0.163	ρ = -0.07 p = 0.695 1-β = 0.066
	Females	ρ = 0.287 p = 0.184 1-β = 0.246	ρ = -0.188 p = 0.389 1-β = 0.129	ρ = 0.259 p = 0.232 1-β = 0.207	ρ = -0.027 p = 0.901 1-β = 0.052
	Overall	ρ = -0.007 p = 0.96 1-β = 0.05	ρ = -0.138 p = 0.308 1-β = 0.164	ρ = 0.182 p = 0.176 1-β = 0.254	ρ = -0.075 p = 0.577 1-β = 0.082
Efficiency measure	Males		ρ = -0.061 p = 0.732 1-β = 0.062	ρ = 0.132 p = 0.458 1-β = 0.109	ρ = -0.014 p = 0.936 1-β = 0.051
	Females		ρ = -0.341 p = 0.111 1-β = 0.335	ρ = 0.028 p = 0.9 1-β = 0.052	ρ = -0.039 p = 0.861 1-β = 0.053
	Overall		ρ = -0.198 p = 0.139 1-β = 0.293	ρ = 0.072 p = 0.593 1-β = 0.08	ρ = -0.036 p = 0.791 1-β = 0.057

Spearman's correlation coefficients (ρ), p values and 1-β values for analysis of the relationship between left hand 2D:4D and standard age, stanine and efficiency measure scores on all subtasks of the Dyscalculia Screener in males (n =34), females (n =24) and the entire data set (n = 58), significant effect indicated in bold.

		Simple RT	**Dot enumeration**	**Number comparison**	**Addition**
Standard age	Males	ρ = -0.219 p = 0.214 1-β = 0.221	ρ = 0.038 p = 0.831 1-β = 0.055	ρ = -0.014 p = 0.938 1-β = 0.051	ρ = 0.104 p = 0.56 1-β = 0.086
	Females	ρ = 0.079 p = 0.72 1-β = 0.063	ρ = -0.161 p = 0.463 1-β = 0.107	ρ = -0.068 p = 0.757 1-β = 0.06	ρ = -0.075 p = 0.732 1-β = 0.062
	Overall	ρ = -0.127 p = 0.346 1-β = 0.146	ρ = -0.08 p = 0.552 1-β = 0.087	ρ = -0.061 p = 0.65 1-β = 0.071	ρ = 0.013 p = 0.923 1-β = 0.051
Stanine	Males	ρ = -0.185 p = 0.294 1-β = 0.169	ρ = -0.089 p = 0.616 1-β = 0.076	ρ = -0.009 p = 0.96 1-β = 0.05	ρ = -0.07 p = 0.695 1-β = 0.066
	Females	ρ = 0.056 p = 0.8 1-β = 0.057	ρ = -0.06 p = 0.784 1-β = 0.058	ρ = -0.054 p = 0.806 1-β = 0.056	ρ = 0.161 p = 0.462 1-β = 0.107
	Overall	ρ = -0.116 p = 0.391 1-β = 0.13	ρ = -0.106 p = 0.432 1-β = 0.116	ρ = -0.074 p = 0.584 1-β = 0.082	ρ = 0.026 p = 0.845 1-β = 0.054
Efficiency measure	Males		ρ = 0.029 p = 0.871 1-β = 0.053	ρ = 0.018 p = 0.921 1-β = 0.051	ρ = 0.112 p = 0.527 1-β = 0.092
	Females		ρ = -0.007 p = 0.975 1-β = 0.05	ρ = -0.064 p = 0.771 1-β = 0.059	ρ = -0.063 p = 0.774 1-β = 0.058
	Overall		ρ = -0.18 p = 0.893 1-β = 0.249	ρ = -0.027 p = 0.844 1-β = 0.054	ρ = 0.023 p = 0.865 1-β = 0.053

Appendix 7

Additional statistical information relating Chapter 9

Section 9.3

Results from the Kolmogorov-Smirnov analysis to explore normality.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RH2D4D	.058	116	.200*	.989	116	.465
LH2D4D	.076	116	.094	.980	116	.082
ZSATlit	.153	116	.000	.898	116	.000
ZSATnum	.170	116	.000	.893	116	.000

*. This is a lower bound of the true significance.

Appendix 8

BASIC NUMERICAL ASSESSMENT TEST BATTERY

FOR CHILDREN 5-7 YEARS OF AGE

Evaluation and Protocol

Participant Number……………………………………………

Sex……………………………………..…………………….

Date…………………………………..……………………...

<u>COUNTING</u>

Task booklet

Instruction: *"I'm going to ask you to do some counting for me. First I will ask you to count some black circles for me and then some pictures of different objects. While you are counting please point to the circles/objects and count aloud. I will then ask you how many circles/objects there are."*

Example

One point for a correctly recited verbal sequence corresponding with synchronised pointing behaviour in which each item was acknowledged once and only once. A second point for the correct identification of the sets' numerical value.

Test Item	**Correct answer**	**Verbal (sequence) error?**	**Verbal number give correct/**	**Score**
1	12			
2	10			
3	7			
4	5			
5	9			
		Total		

READING NUMBERS

Task booklet

Instruction: *"I'm going to show you some numbers. Please tell me what the numbers are."*

Example

Test Item	**Stimulus**	**Response correct?**	**Score**
1	2		
2	6		
3	8		
4	9		
5	13		
6	15		
7	17		
8	20		
9	23		
10	25		
		Total	

<u>WRITING NUMBERS</u>

Work booklet

Instruction: *"I'm going to ask you to write down some numbers for me in your workbook."*

Example *"Please can you write the number 1 on the first line."*

Two points for a correct response, one point for a correct response following repetition.

Test Item	**Stimulus**	**Response correct?**	**Repetition?**	**Score**
1	3			
2	9			
3	5			
4	10			
5	14			
		Total		

<u>MENTAL ARITHMETIC</u>

Instruction: *"I'm going to read out some number for you to either add or subtract in your head."*

Two points for a correct response, one point for a correct response following repetition.

"Please can you add together the following numbers in your head and tell me the answer."

Test Item	Stimulus	Response correct?	Repetition?	Score
1	1 + 3 (4)			
2	4 + 2 (6)			
3	6 + 5 (11)			
		Total		

"Please can you subtract the following numbers in your head and tell me the answer."

Test Item	Stimulus	Response correct?	Repetition?	Score
1	10 - 5 (5)			
2	6 - 4 (2)			
3	9 - 7 (2)			
		Total		

<u>MENTAL ARITHMETIC</u>

Instruction: *"I'm now going to read to you some sentences. Each sentence will be followed by a question please answer the question as best you can."*

Two points for a correct response, one point for a correct response following repetition

Test Item	Stimulus	Response correct?	Repetition?	Score
1	There are three people on the bus. One more person gets on how many are now on the bus? **(3+1=4)**			
2	There are three birds in a nest. They are joined by two more birds. How many birds are now in the nest? **(3+2=5)**			
3	There are five books on the shelf. Three are taken away. How many books are now on the shelf? **(5-3=2)**			
4	There are ten children in the playground. Two children leave. How many children are now in the playground? **(10-2=8)**			
		Total		

NUMBER LINE TASKS

Task booklet

Instruction: *"I'm going to show you some number lines. Below each line is an arrow. I'm going to ask you what number you think the arrow is pointing to."*

Example

Test Item	Stimulus	Response	Score
1	1		
2	7		
		Total	

Two points for a correct response, one point for a response correct within 1 number either side

Test Item	Stimulus	Response	Score
1	8		
2	3		
		Total	

NUMBER LINE TASKS

Work booklet

Instruction: *"We are now going to look at some more number lines in your workbook. I'm going to tell you a number and ask you to make a mark where you think that number should go on the number line."*

Example "*Please can you mark where you think the number five should go on the top line in your workbook"*

Two points for a correct response, one point for a response correct within 1 number either side

Test Item	Stimulus	Response	Score
1	9		
2	2		
		Total	

NUMBER LINE TASKS

Task booklet

Instruction: *"I'm now going to show you some rectangles. Some of them have numbers inside of them and some do not. I'm going tell you a number and ask you to point to the empty rectangle which you think the number should be in."*

Example "*Please can you point to the empty rectangle which the number four belongs."*

Test Item	Stimulus	Response correct?	Score
1	6		
2	8		
		Total	

NUMBER LINE TASKS

Task booklet and Work booklet

Instruction: *"I'm going to show you some lines of numbers and ask you to write the numbers out in your workbook from the smallest to the biggest."*

Example

Test Item	Correct Answer	Response correct?	Score
1	2 3 4 5 6 7		
2	2 4 5 7 9 10		
		Total	

NUMBER LINE TASKS

Task booklet

Instruction: *"I'm going to show you some lines of pictures and ask you to point to a particular number picture in the line"*

Example *"Please can you point to the second cat in the line."*

Test Item	Stimulus	Response correct?	Score
1	4^{th} teddy		
2	6^{th} wizard		
		Total	

NUMBER COMPARISON

Task booklet

Instruction: *"I'm going to show you two circles each with a different number of dots inside and ask you to, as quickly as possible, point to the circle with the most dots inside."*

Example

Test Item	Correct answer	Response correct?	Score
1	RIGHT		
2	LEFT		
3	RIGHT		
4	RIGHT		
5	LEFT		
		Total	

<u>NUMBER COMPARISON</u>

Task booklet

Instruction: *"You will now see two circles, this time one will have a number inside and one will have dots inside. I'm going to ask you, again as quickly as possible, to tell me whether the number of dots in one circle is the same as the number in the other circle"*

Example

Test Item	Correct answer	Response correct?	Score
1	NO		
2	YES		
3	NO		
4	YES		
5	YES		
		Total	

<u>NUMBER COMPARISON</u>

Work booklet

Instruction: *"You will now see some pairs of numbers in your workbook. For each pair of numbers please circle the biggest. "*

Example

Test Item	Stimulus		Response correct?	Score
1	5	**15**		
2	13	**31**		
3	79	**81**		
4	**96**	69		
			Total	

ESTIMATION

Work booklet

Instruction: *"I'm going to show you some different pictures. For each page I'm going to ask you to tell me how many pictures you think you can see. There will be too many for you to count so please guess how many you think there are."*

Test Item	Stimulus	Response	Score*
1	Dogs (17)		
2	Parrots (28)		
3	Trees (38)		
4	Balls (57)		
5	Cups (89)		
		Total	

***Scoring**

Item 1
2 points if between 15 and 20
1 point if between 12 and 23

Item 2
2 points if between 25 and 32
1 point if between 22 and 35

Item 3
2 points if between 35 and 42
1 point if between 32 and 45

Item 4
2 points if between 45 and 70
1 point if between 25 and 80

Item 5
2 points if between 70 and 110
1 point if between 35 and 125

		Score	Max
1	Counting		10
2	Reading numbers		10
3	Writing numbers		10
4	Mental arithmetic (addition)		6
5	Mental arithmetic (subtraction)		6
6	Mental arithmetic (everyday)		8
7	Number line tasks (exact identification)		2
8	Number line tasks (approx. identification)		4
9	Number line tasks (approx positioning)		4
10	Number line tasks (identifying the missing number)		2
11	Number line tasks (arranging)		2
12	Number line tasks (positioning)		2
13	Number comparison (dots)		5
14	Number comparison (dots vs. Arabic number)		5
15	Number comparison (Arabic numbers)		4
16	Estimation		10
	Total		90

		Score	Max
1	Counting		10
2	Reading numbers		10
3	Writing numbers		10
4	Mental arithmetic		20
5	Number line tasks		16
6	Number comparison		14
7	Estimation		10
	Total		90

Appendix 9

BASIC
NUMERICAL
ASSESSMENT
TEST BATTERY

FOR CHILDREN
8-11 YEARS OF AGE

Evaluation and Protocol

Participant Number……………………………………………

Sex…………………………………….……………………….

Date………………………………….………………………...

COUNTING

Task booklet

Instruction: *"I'm going to ask you to do some counting for me. First I will ask you to count some black circles for me and then some pictures of different objects. While you are counting please point to the circles/objects and count aloud. I will then ask you how many circles/objects there are."*

Example

One point for a correctly recited verbal sequence corresponding with synchronised pointing behaviour in which each item was acknowledged once and only once. A second point for the correct identification of the sets' numerical value.

Test Item	Correct answer	Verbal (sequence) error?	Verbal number give correct/	Score
1	12			
2	10			
3	18			
4	7			
5	5			
6	9			
7	14			
8	11			
9	17			
10	15			
		Total		

READING NUMBERS

Task booklet

Instruction: *"I'm going to show you some numbers. Please tell me what the numbers are."*

Example

Test Item	Stimulus	Response correct?	Score
1	2		
2	6		
3	8		
4	13		
5	15		
6	17		
7	20		
8	23		
9	31		
10	50		
11	138		
12	305		
13	785		
14	1179		
15	1900		
16	3002		
17	6485		
18	8057		
19	18000		
20	63002		
		Total	

<u>WRITING NUMBERS</u>

Work booklet

Instruction: *"I'm going to ask you to write down some numbers for me in your workbook."*

Example *"Please can you write the number 1 on the first line."*

Two points for a correct response, one point for a correct response following repetition.

Test Item	Stimulus	Response correct?	Repetition?	Score
1	3			
2	9			
3	5			
4	10			
5	14			
6	38			
7	1200			
8	503			
9	169			
10	4658			
		Total		

MENTAL ARITHMETIC

Instruction: *"I'm going to read out some number for you to either add or subtract in your head."*

Two points for a correct response, one point for a correct response following repetition.

"Please can you add together the following numbers in your head and tell me the answer."

Test Item	Stimulus	Response correct?	Repetition?	Score
1	6 + 5 (11)			
2	12 + 6 (18)			
3	13 + 19 (32)			
4	15 + 12 (27)			
5	26 + 22 (48)			
		Total		

"Please can you subtract the following numbers in your head and tell me the answer."

Test Item	Stimulus	Response correct?	Repetition?	Score
1	9 – 7 (2)			
2	19 – 6 (13)			
3	27 - 17 (10)			
4	25 - 12 (13)			
5	34 - 26 (8)			
		Total		

MENTAL ARITHMETIC

"I'm now going to read to you some sentences. Each sentence will be followed by a question please answer the question as best you can."

Two points for a correct response, one point for a correct response following repetition.

Test Item	Stimulus	Response correct?	Repetition?	Score
1	There are ten children in the playground. Two children leave. How many children are now in the playground? **(10-2=8)**			
2	The are 26 children in class 1. There are 10 more children in class 1 than class 2. How many children are in class 2? **(26-10=16)**			
3	Peter has sixteen marbles. He has four less than Anne how many marbles does Anne have? **(16+4=20)**			
4	Sarah has 12 apples. She gives 5 apples to Mark. How many apples does she have left? **(12-5=7)**			
5	There are 76 books on the shelf. I remover 50 books. How many are now on the shelf? **(76-50=26)**			
		Total		

NUMBER LINE TASKS

Task booklet

Instruction: *"I'm going to show you some number lines. Below each line is an arrow. I'm going to ask you what number you think the arrow is pointing to."*

Examples

Test Item	Stimulus	Response	Score
1	7		
2	40		
3	80		
4	600		
		Total	

Two points for a correct response, one point for a response correct within 1 number either side

Test Item	Stimulus	Response	Score
1	8		
2	3		
3	7		
4	4		
		Total	

NUMBER LINE TASKS

Work booklet

Instruction: *"We are now going to look at some more number lines in your workbook. I'm going to tell you a number and ask you to make a mark where you think that number should go on the number line."*

Example "*Please can you mark where you think the number five should go on the top line in your workbook."*

Two points for a correct response, one point for a response correct within 1 number either side

Test Item	Stimulus	Response	Score
1	9		
2	2		
3	50		
4	10		
		Total	

NUMBER LINE TASKS

Task booklet

Instruction: *"I'm now going to show you some rectangles. Some of them have numbers inside of them and some do not. I'm going tell you a number and ask you to point to the empty rectangle which you think the number should be in."*

Example "*Please can you point to the empty rectangle which the number four belongs."*

Test Item	Stimulus	Response correct?	Score
1	8		
2	36		
3	700		
4	140		
		Total	

<u>NUMBER LINE TASKS</u>

Task booklet and Work booklet

Instruction: *"I'm going to show you some lines of numbers and ask you to write the numbers out in your workbook from the smallest to the biggest."*

Example

Test Item	Correct Answer	Response correct?	Score
1	2 4 5 7 9 10		
2	6 7 12 15 17 22		
3	99 121 159 179 212 247		
4	2154 2451 4521 5124 5214		
		Total	

<u>NUMBER LINE TASKS</u>

Task booklet

Instruction: *"I'm going to show you some lines of pictures and ask you to point to a particular number picture in the line"*

Example *"Please can you point to the second cat in the line."*

Test Item	Stimulus	Response correct?	Score
1	4^{th} teddy		
2	6^{th} wizard		
3	5^{th} sunflower		
4	9^{th} star		
		Total	

NUMBER COMPARISON

Task booklet

Instruction: *"I'm going to show you two circles each with a different number of dots inside and ask you to, as quickly as possible, point to the circle with the most dots inside."*

Example

Test Item	Correct answer	Response correct?	Score
1	LEFT		
2	RIGHT		
3	RIGHT		
4	RIGHT		
5	LEFT		
6	RIGHT		
7	LEFT		
8	RIGHT		
9	LEFT		
10	RIGHT		
		Total	

NUMBER COMPARISON

Task booklet

Instruction: *"You will now see two circles, this time one will have a number inside and one will have dots inside. I'm going to ask you, again as quickly as possible, to tell me whether the number of dots in one circle is the same as the number in the other circle"*

Example

Test Item	Correct answer	Response correct?	Score
1	NO		
2	YES		
3	NO		
4	YES		
5	YES		
6	NO		
7	YES		
8	YES		
9	NO		
10	YES		
		Total	

NUMBER COMPARISON

Work booklet

Instruction: *"You will now see some pairs of numbers in your workbook. For each pair of numbers please circle the biggest. "*

Example

Test Item	Stimulus		Response correct?	Score
1	13	**31**		
2	79	**81**		
3	1007	**1070**		
4	**511**	298		
5	**654**	546		
6	9768	**35201**		
7	**96**	69		
8	**201**	102		
			Total	

ESTIMATION

Work booklet

Instruction: *"I'm going to show you some different pictures. For each page I'm going to ask you to tell me how many pictures you think you can see. There will be too many for you to count so please guess how many you think there are."*

Test Item	Stimulus	Response	Score*
1	Dogs (17)		
2	Parrots (28)		
3	Trees (38)		
4	Balls (57)		
5	Cups (89)		
		Total	

***Scoring**

Item 1
2 points if between 15 and 20
1 point if between 12 and 23

Item 2
2 points if between 25 and 32
1 point if between 22 and 35

Item 3
2 points if between 35 and 42
1 point if between 32 and 45

Item 4
2 points if between 45 and 70
1 point if between 25 and 80

Item 5
2 points if between 70 and 110
1 point if between 35 and 125

		Score	Max
1	Counting		20
2	Reading numbers		20
3	Writing numbers		20
4	Mental arithmetic (addition)		10
5	Mental arithmetic (subtraction)		10
6	Mental arithmetic (everyday)		10
7	Number line tasks (exact identification)		4
8	Number line tasks (approx. identification)		8
9	Number line tasks (approx positioning)		8
10	Number line tasks (identifying the missing number)		4
11	Number line tasks (arranging)		4
12	Number line tasks (positioning)		4
13	Number comparison (dots)		10
14	Number comparison (dots vs. Arabic number)		10
15	Number comparison (Arabic numbers)		8
16	Estimation		10
	Total		160

		Score	Max
1	Counting		20
2	Reading numbers		20
3	Writing numbers		20
4	Mental arithmetic		30
5	Number line tasks		32
6	Number comparison		28
7	Estimation		10
	Total		160

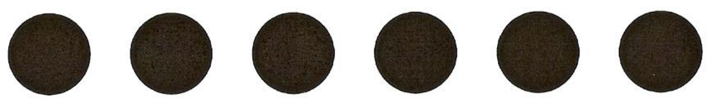
Appendix 10

Example visual stimuli for the basic numerical test battery

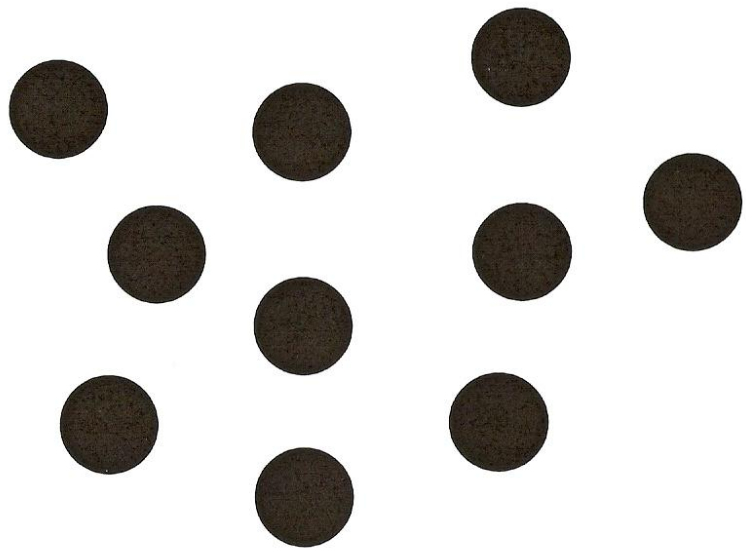
The following images are examples of some of the stimuli used in tasks which required visual presentation of material. For presentation purposes the size of some of the stimuli has been reduced relative to how it appeared in the numerical test battery.

Counting

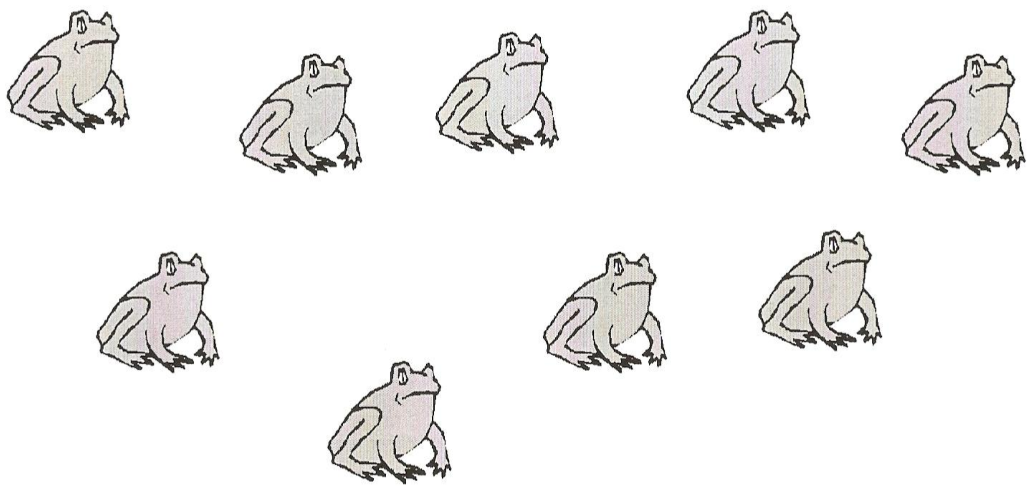
Example 1



Example 2



Example 3

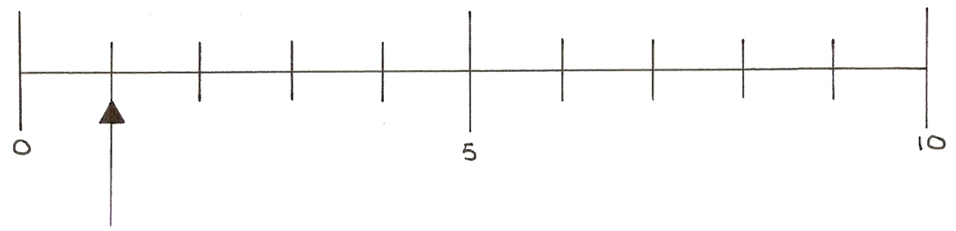


Reading numbers

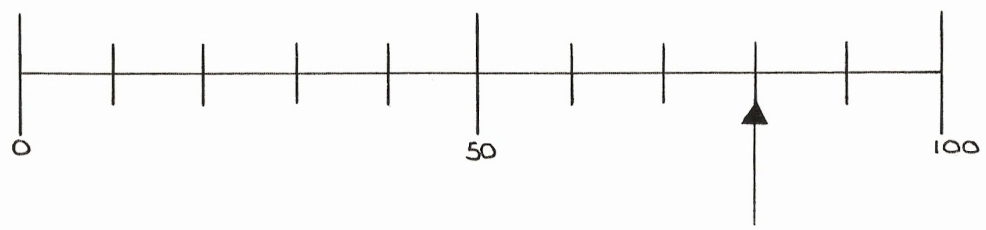
12

Identification of number on an analogue scale

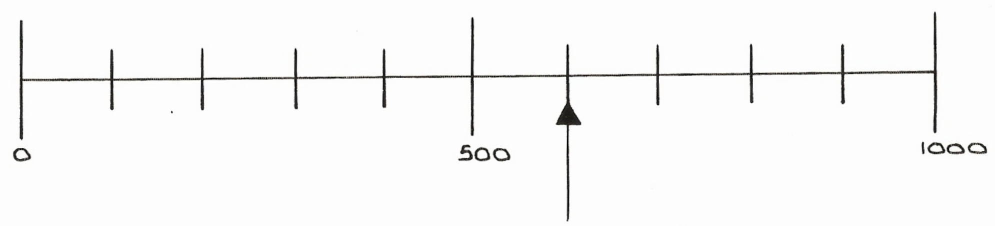
Example 1



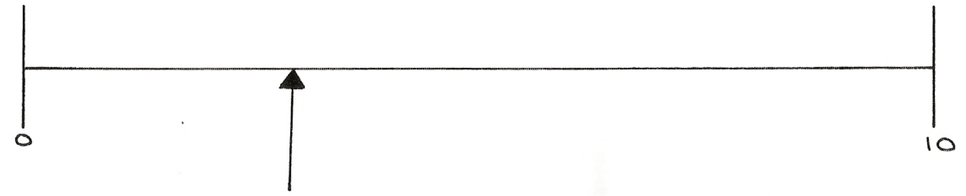
Example 2



Example 3

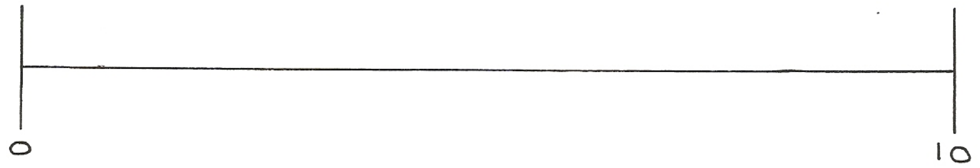


Approximate identification of number on an analogue scale



Approximate positioning of number on an analogue scale

Example 1



Example 2

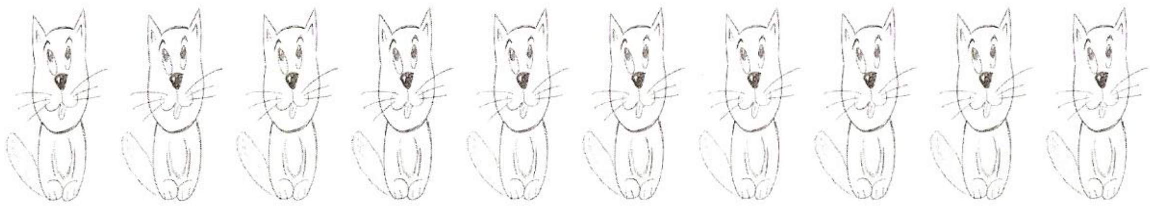
Identification of the missing number in a numerical series

5		7		9		11	12

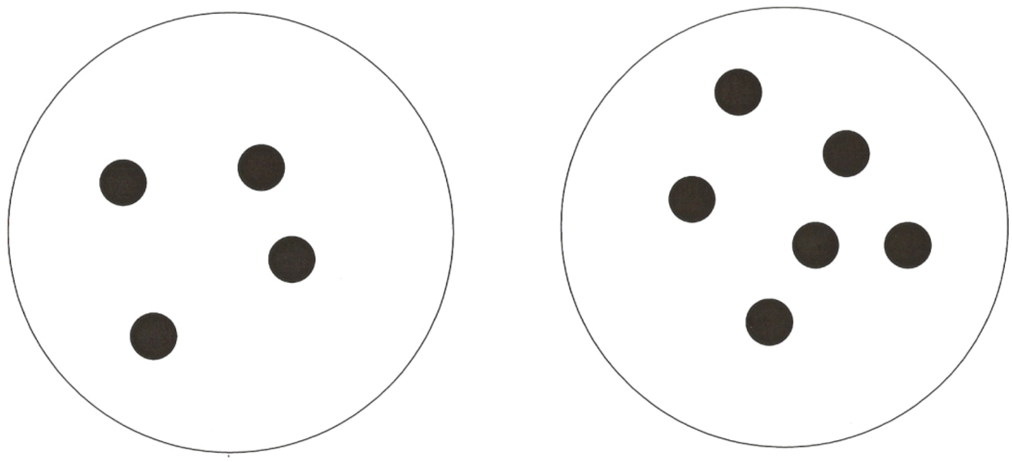
Arranging numbers

5 7 9 10 4 2

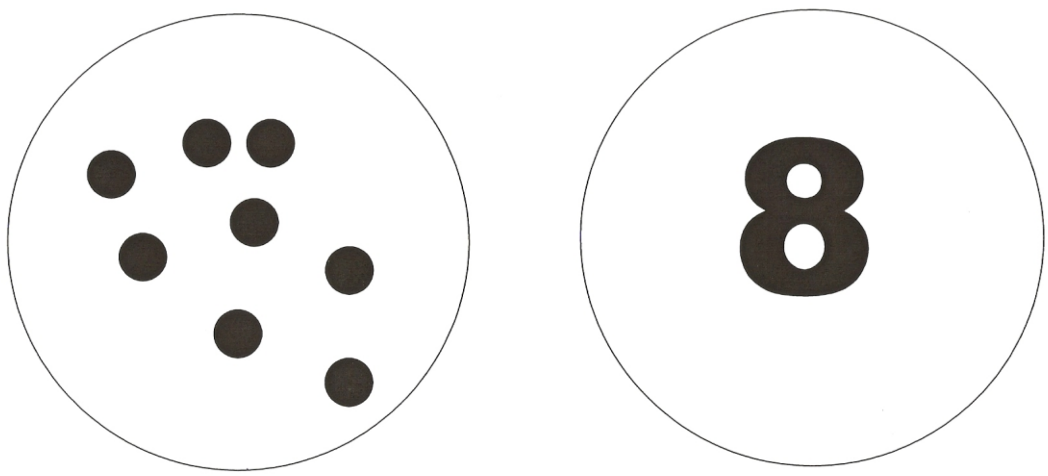
Positioning numbers



Quantity comparison



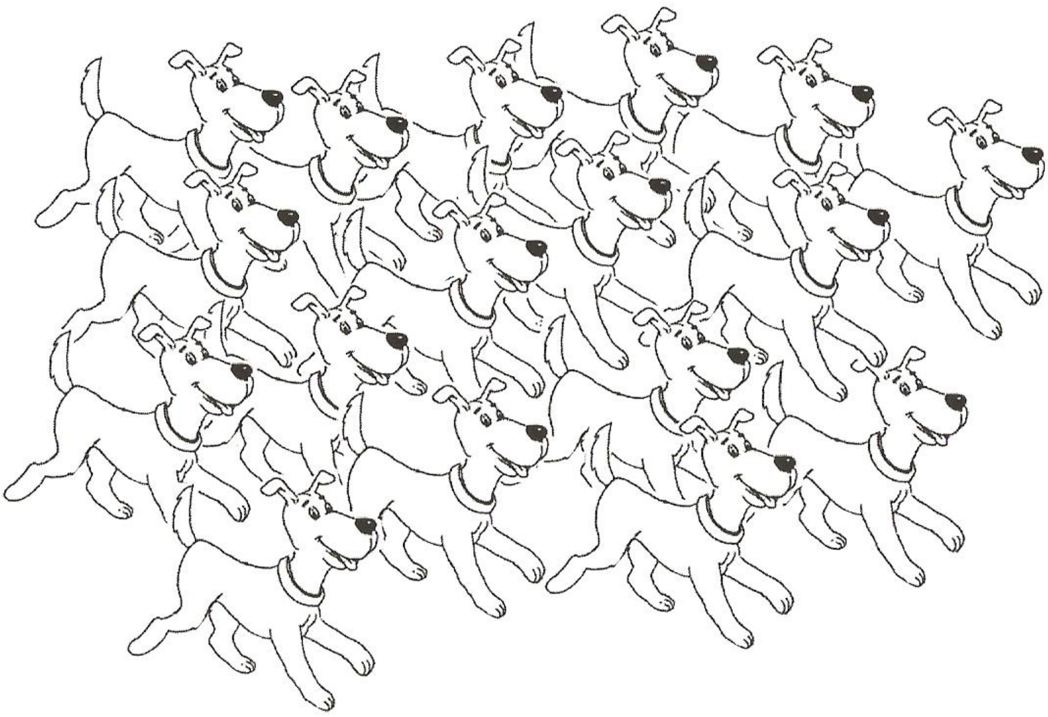
Arabic digit-quantity comparison



Arabic digit comparison

13……………………………………………….31

Estimation



Appendix 11

Descriptive statistics relating to performance on the numerical test batteries

Task	5-7 years old (n = 41)				8-11 years old (n = 32)		
	Mean	Min-Max	SD		Mean	Min-Max	SD
Counting	96.83	50-100	9.34		97.03	80-100	5.51
Reading numbers	91.46	10-100	19.69		91.41	55-100	10.57
Writing numbers	89.76	0-100	19.56		95.63	70-100	9.22
Addition	73.17	0-100	30.25		81.88	10-100	22.06
Subtraction	58.13	0-100	30.76		60.31	0-100	27.88
Everyday addition and subtraction	72.87	0-100	27.37		65.31	20-100	20.48
Mental Arithmetic overall	68.54	15-100	24.53		68.75	13.33-93.33	18.47
Identification of a number on an analogue scale	87.8	0-100	24.45		91.41	50-100	16.33
Approximate identification of number on an analogue scale	49.39	0-100	28.22		78.13	25-100	22.67
Approximate positioning of number on analogue scale	47.56	0-75	21.51		58.59	25-87.5	15.37
Identification of the missing number/s in a numerical series	96.34	50-100	13.18		94.06	10-100	17.48
Arranging numbers	73.17	0-100	37.25		89.06	25-100	17.89
Positioning numbers	92.68	0-100	21.1		99.22	75-100	4.42
Number line tasks overall	68.45	25-93.75	14.11		79.46	3.13-100	17.17
Quantity comparison	87.07	10-100	19.27		91.88	50-100	13.30
Arabic digit-quantity comparison	73.17	40-100	19.8		88.59	60-100	11.45
Arabic digit comparison	80.49	0-100	24.69		90.16	20-100	20.56
Number comparison overall	81.88	50-100	13.66		91.96	75-100	7.09
Estimation	23.66	0-100	21.3		39.68	10-100	22.36
Test battery overall	73.69	46.67-90	11.13		83.32	52.5-93.13	8.42

Appendix 12

Edinburgh Handedness Inventory

Please indicate your hand preferences in the following activities *by putting a check in the appropriate column*. Where the preference is so strong that you would never try to use the other hand, unless absolutely forced to, *put 2 checks*. If in any case you are really indifferent, *put a check in both columns*.

Some of the activities listed below require the use of both hands. In these cases, the part of the task, or object, for which hand preference is wanted is indicated in brackets.

Please try and answer all of the questions, and only leave a blank if you have no experience at all with the object or task.

	Left	Right
1. Writing	☐ ☐	☐ ☐
2. Drawing	☐ ☐	☐ ☐
3. Throwing	☐ ☐	☐ ☐
4. Scissors	☐ ☐	☐ ☐
5. Toothbrush	☐ ☐	☐ ☐
6. Knife (without fork)	☐ ☐	☐ ☐
7. Spoon	☐ ☐	☐ ☐
8. Broom (upper hand)	☐ ☐	☐ ☐
9. Striking Match (match)	☐ ☐	☐ ☐
10. Opening box (lid)	☐ ☐	☐ ☐
TOTAL		

(to be completed by the experimenter)

Difference	Cumulative TOTAL	Result

Appendix 13

SNARC analysis according to the procedure followed by Bull and Benson (2006).

According to the procedure followed by Bull and Benson (2006) analysis of the presence of a SNARC effect (described in chapter 6, section 6.4.4) was repeated according to digit ratio and sex. Results of these analyses are summarised in the table below.

Regression analysis of SNARC effect for low and high 2D:4D groups on the SNARC and Vowel-consonant control task.

	Regression weights compared to 0			
	SNARC		Vowel-Consonant	
	t	p	*t*	p
Overall (*n = 67*)	**4.389**	**< 0.001**	1.501	0.147
Males (*n = 35*)	**3.196**	**0.003**	**2.064**	**0.047**
Females (*n = 32*)	**2.96**	**0.006**	0.23	0.82
Right Hand				
Low 2D:4D (*n = 33*)	**3.035**	**0.005**	0.243	0.809
High 2D:4D (*n = 34*)	**3.138**	**0.004**	**2.624**	**0.013**
Low 2D:4D males (*n = 17*)	**2.326**	**0.033**	0.073	0.942
High 2D:4D males (*n = 18*)	**2.133**	**0.048**	**3.423**	**0.003**
Low 2D:4D females (*n = 16*)	1.915	0.075	0.377	0.711
High 2D:4D females (*n = 16*)	**2.276**	**0.038**	0.767	0.455
Left Hand				
Low 2D:4D (*n = 34*)	**4.045**	**< 0.001**	0.721	0.477
High 2D:4D (*n = 34*)	**2.158**	**0.038**	1.45	0.156
Low 2D:4D males (*n = 16*)	**3.569**	**0.003**	1.723	0.104
High 2D:4D males (*n = 19*)	1.189	0.251	1.124	0.277
Low 2D:4D females (*n = 18*)	**2.219**	**0.044**	0.577	0.573
High 2D:4D females (*n = 15*)	1.994	0.063	0.939	0.362

In the separate analysis of male and female data regression weights differed significantly from 0 in both sexes indicating the presence of a significant SNARC effect for both groups. Similarly the separate analysis of participants classified as high 2D:4D vs. low 2D:4D on the basis of both right and left hand median values revealed the presence of a significant SNARC effect in both sets of participants with regression weights from high and low participants observed to differ significantly from 0. Analysis according to both sex and right hand digit ratio group revealed a significant SNARC effect in low and high 2D:4D males and high 2D:4D females. When analysis was split according to both sex and left hand digit ratio however only low 2D:4D (high T) group males and low 2D:4D females demonstrated SNARC regression weightings that were revealed to be significantly greater than 0. A significant SNARC effect was not revealed for high left hand 2D:4D (low T) males or females.

With regard to the control vowel-consonant task, in line with evidence for an ordinal representation of letters of the alphabet, regression weightings for male data analysed independently were revealed to be significantly different from 0. Female regression weightings however were not found to be significantly different from 0. In analysis of the vowel-consonant control task repeated according to right hand digit ratio group and sex, significant effects were revealed for high right hand 2D:4D (low testosterone) participants overall (male and female data combined) and for high right hand 2D:4D males. Regression weighting for all other groups however were found not to be significantly different from 0. Analysis of findings from the vowel-consonant control task according to left hand digit ratio group and sex showed no significant effects.

Bull and Benson (2006) also considered the simple presence of a negative slope (indicating faster left hand responses for lower magnitudes/letters at the beginning of the alphabet), irrespective of steepness of slope and differences from 0. Mirroring such analysis in the current study, the percentage of each group showing a negative slope on the SNARC task is presented in the table below.

Percentage of participants showing a negative slope and chi squared analysis of group differences in slope direction for high and low 2D:4D, male and female participants.

	Group comparisons SNARC task			Group comparisons Vowel-consonant task		
	% showing negative slope	χ^2	p	% showing negative slope	χ^2	p
Overall (*n = 67*)	73.1			56.7		
Males (*n = 35*)	77.1			62.9		
Females (*n = 32*)	68.8	0.599	0.439	50	1.126	0.289
Righ Hand						
Low 2D:4D (*n = 33*)	75.8			45.5		
High 2D:4D (*n = 34*)	70.6	0.228	0.633	67.6	3.36	0.067
Low 2D:4D males (*n = 17*)	82.4			47.1		
High 2D:4D males (*n = 18*)	72.2	0.509	0.476	77.8	3.534	0.06
Low 2D:4D females (*n = 16*)	68.8			43.8		
High 2D:4D females (*n = 16*)	68.8	0	1	56.3	0.5	0.48
Left Hand						
Low 2D:4D (*n = 32*)	78.1			56.3		
High 2D:4D (*n = 35*)	68.6	0.777	0.378	57.1	0.005	0.941
Low 2D:4D males (*n = 17*)	82.4			64.7		
High 2D:4D males (*n = 18*)	72.2	0.509	0.476	61.1	0.048	0.826
Low 2D:4D females (*n = 15*)	73.3			46.7		
High 2D:4D females (*n = 17*)	64.7	0.276	0.599	52.9	0.125	0.723

As can be seen in the above table, for SNARC task analysis, for all comparisons excluding than between high and low right hand 2D:4D females, low 2D:4D participants showed a consistently higher frequency of negative slopes, the percentage of low and high 2D:4D participants demonstrating negative slopes however was largely very similar across all comparisons. Chi squared analysis revealed no significant differences in slope direction between any digit ratio group comparisons on the SNARC task. A contradictory pattern of results was apparent for data derived from the vowel-consonant control task where high 2D:4D participants (categorised according to both left and right 2D:4D measures) showed a higher frequency of negative slopes than low 2D:4D participants in all comparisons excluding that between low and high left hand 2D:4D males. Again however Chi squared analysis revealed no significant differences in slope direction between any digit ratio group comparisons on the vowel-consonant task. For both the SNARC and vowel-consonant tasks, males showed a higher percentage of negative slopes as compared to females, chi squared analysis however revealed no significant sex differences on either task.

Appendix 14

Edinburgh Handedness Inventory (Adapted)

Please indicate your hand preferences in the following activities *by putting a check in the appropriate column*. Where the preference is so strong that you would never try to use the other hand, unless absolutely forced to, *put 2 checks*. If in any case you are really indifferent, *put a check in both columns*.

Some of the activities listed below require the use of both hands. In these cases, the part of the task, or object, for which hand preference is wanted is indicated in brackets.

Please try and answer all of the questions, and only leave a blank if you have no experience at all with the object or task.

	Left	Right
1. Writing	☐ ☐	☐ ☐
2. Drawing	☐ ☐	☐ ☐
3. Throwing	☐ ☐	☐ ☐
4. Scissors	☐ ☐	☐ ☐
5. Toothbrush	☐ ☐	☐ ☐
6. Knife (without fork)	☐ ☐	☐ ☐
7. Spoon	☐ ☐	☐ ☐
8. Using TV Remote	☐ ☐	☐ ☐
9. Comb Hair	☐ ☐	☐ ☐
10. Holding can or bottle to open it	☐ ☐	☐ ☐
TOTAL		

(to be completed by the experimenter)

Difference	Cumulative TOTAL	Result

Appendix 15

Analysis of the relationship between KS1 SAT numeracy and literacy scores according to the procedure followed by Brosnan (2008)

In addition to the data reported in chapter 9 Brosnan (2008) also reports, average 2D:4D (calculated as the mean of left and right hand 2D:4D measures), variability between left and right hand 2D:4D (calculated as right hand 2D:4D – left hand 2D:4D and notated as Dr-1) and a Z-SAT difference score (calculated as Z-SAT numeracy minus Z-SAT literacy, thus a positive score indicates a high numeracy score relative to literacy). In line with the analysis conducted by Brosnan (2008), the table below displays the results of t-test analysis to explore sex differences in the additional measures of 2D:4D reported by Brosnan (2008) and in Z-SAT scores.

Average 2D:4D values and t-test analysis of sex differences in average 2D:4D, Dr-1 and Z-SAT scores (significant effect highlighted in bold).

Measure	Males	Females	Sex differences		
	Mean (SD)	Mean (SD)	t	df	p
Average 2D:4D	0.951 (0.029)	0.961 (0.03)	1.794	114	0.075
Dr-1	0.007 (0.029)	0.007 (0.03)	0.086	114	0.932
Z-SAT numeracy	0.056 (0.951)	-0.061 (1.056)	0.636	117	0.526
Z-SAT literacy	-0.072 (1.029)	0.078 (0.97)	0.816	117	0.416
Z-SAT (numeracy- literacy)	0.128 (0.769)	-0.139 (0.689)	**1.988**	**117**	**0.049**

As can be seen in the table above, no significant sex differences were revealed for average digit ratio or the difference between left and right hand digit ratio (Dr – 1). No significant sex difference were found for analysis of Z-SAT numeracy and literacy scores, a significant sex difference however was revealed for analysis of the difference between numeracy and literacy (calculated as Z-SAT numeracy – Z-SAT literacy). In males a positive Z-SAT difference score was revealed suggesting improved numeracy performance relative to literacy, in contrast a negative Z-SAT difference score was revealed in females suggesting improved literacy performance relative to numeracy.

The tables below display the relationship between 2D:4D measures and Z-SAT scores for analysis of the entire sample and male and female data considered separately. Analysis of all data revealed a significant positive correlation between the difference between left and right hand digit ratio (Dr-1) and Z-SAT literacy scores thus higher 2D:4D on the right hand as compared to the left (thought to indicate lower prenatal testosterone exposure) was associated with higher Z-SAT literacy scores. No further association between 2D:4D and Z-SAT scores were found to be significant in analysis of the entire sample. No significant association were revealed in males. In females a significant positive correlation was identified between right hand 2D:4D and Z-SAT literacy scores suggesting higher SAT literacy scores in females with higher right hand 2D:4D values. Analysis of female data also revealed a significant positive correlation between Dr-1 and SAT numeracy and literacy scores suggesting that higher 2D:4D on the right hand as compared to the left may be associated with higher Z-SAT literacy and numeracy scores in females.

Pearson's correlation coefficients and associated p values of associations between measures of 2D:4D and Z-SAT scores in all data (i.e. male and female data combined), significant effect indicated in bold.

		Z-SAT numeracy	**Z-SAT literacy**	**Z-SAT (numeracy-literacy)**
2D:4D	**Right hand**	r = 0.087, p = 0.348	r = 0.16, p = 0.085	r = 0.079, p = 0.539
	Left hand	r = -0.027, p = 0.776	r = 0.005, p = 0.96	r = 0.083, p = 0.519
	Average	r = 0.038, p = 0.683	r = 0.098, p = 0.297	r = 0.091, p = 0.483
	Dr-1	r = 0.151, p = 0.106	**r = 0.203, p = 0.029**	r = -0.001, p = 0.995

Pearson's correlation coefficients and associated p values of associations between measures of 2D:4D and Z-SAT scores in males.

		Z-SAT numeracy	Z-SAT literacy	Z-SAT (numeracy-literacy)
2D:4D	**Right hand**	r = -0.01, p = 0.936	r = -0.069, p = 0.594	r = 0.079, p = 0.539
	Left hand	r = 0.008, p = 0.952	r = -0.055, p = 0.671	r = 0.083, p = 0.519
	Average	r = -0.002, p = 0.99	r = -0.069, p = 0.592	r = 0.09, p = 0.483
	Dr-1	r = -0.02, p = 0.874	r = -0.018, p = 0.888	r = -0.001, p = 0.995

Pearson's correlation coefficients and associated p values of associations between measures of 2D:4D and Z-SAT scores in females (significant effects indicated in bold).

		Z-SAT numeracy	Z-SAT literacy	Z-SAT (numeracy-literacy)
2D:4D	**Right hand**	r = 0.196, p = 0.152	**r = 0.319, p = 0.003**	r = -0.248, p = 0.068
	Left hand	r = -0.052, p = 0.706	r = 0.049, p = 0.721	r = -0.142, p = 0.295
	Average	r = 0.092, p = 0.509	r = 0.266, p = 0.052	r = -0.223, p = 0.106
	Dr-1	**r = 0.328, p = 0.016**	**r = 0.474, p <0.001**	r = -0.154, p = 0.265

Appendix 16

Prospective Power Analysis

Correlational analysis

G*Power Output

Exact – Correlation: Bivariate normal model			
Options:	exact distribution		
Analysis:	A priori: Compute required sample size		
Input:	Tail(s)	=	Two
	Correlation ρ H1	=	0.2
	α err prob	=	0.05
	Power (1–β err prob)	=	0.8
	Correlation ρ H0	=	0
Output:	Lower critical r	=	–0.1412906
	Upper critical r	=	0.1412906
	Total sample size	=	193
	Actual power	=	0.8000846

ANOVA analysis

The sample size requirement for a 2 x 2 x 2 x 2 ANOVA analysis with 2 independent measures variables and 2 repeated measures variables, all with two levels was calculated by computing the sample size needed in order to achieve a small-to-moderate effect size (entered as 0.175 in line with the conversions described by Cohen, 1988) and power of 0.8 for each possible main and interaction effect. As can be seen from the G*Power calculations below the largest required sample size is 259, thus at least 65 participants would be required in each group (260 overall) in order to meet the sample size requirements of each main and interaction effect.

Main effect of an independent measures variable (One factor with 2 levels)

F tests – ANOVA: Fixed effects, special, main effects and interactions			
Analysis:	A priori: Compute required sample size		
Input:	Effect size f	=	0.175
	α err prob	=	0.05
	Power (1–β err prob)	=	0.8
	Numerator df	=	1
	Number of groups	=	2
Output:	Noncentrality parameter λ	=	7.9318750
	Critical F	=	3.8778963
	Denominator df	=	257
	Total sample size	=	259
	Actual power	=	0.8011842

Main effect of a repeated measures variable (One factor with 2 levels)

F tests – ANOVA: Repeated measures, within factors			
Analysis:	A priori: Compute required sample size		
Input:	Effect size f	=	0.175
	α err prob	=	0.05
	Power (1–β err prob)	=	0.8
	Number of groups	=	1
	Number of measurements	=	2
	Corr among rep measures	=	0.5
	Nonsphericity correction ϵ	=	1
Output:	Noncentrality parameter λ	=	8.2075000
	Critical F	=	3.9862695
	Numerator df	=	1.0000000
	Denominator df	=	66.0000000
	Total sample size	=	67
	Actual power	=	0.805929

2 x 2 interaction effect involving 2 independent measures variables (Both one factor with 2 levels)

F tests – ANOVA: Fixed effects, special, main effects and interactions			
Analysis:	A priori: Compute required sample size		
Input:	Effect size f	=	0.175
	α err prob	=	0.05
	Power (1–β err prob)	=	0.8
	Numerator df	=	1
	Number of groups	=	4
Output:	Noncentrality parameter λ	=	7.9318750
	Critical F	=	3.8781841
	Denominator df	=	255
	Total sample size	=	259
	Actual power	=	0.8011611

2 x 2 interaction effect involving 2 repeated measures variables (Both one factor with 2 levels)

F tests – ANOVA: Repeated measures, within factors			
Analysis:	A priori: Compute required sample size		
Input:	Effect size f	=	0.175
	α err prob	=	0.05
	Power (1–β err prob)	=	0.8
	Number of groups	=	1
	Number of measurements	=	4
	Corr among rep measures	=	0.5
	Nonsphericity correction ϵ	=	1
Output:	Noncentrality parameter λ	=	11.2700000
	Critical F	=	2.6716764
	Numerator df	=	3.0000000
	Denominator df	=	135
	Total sample size	=	46
	Actual power	=	0.8018201

2 x 2 interaction effect involving 1 independent measures variable and 1 repeated measures variable (Both one factor with 2 levels)

F tests – ANOVA: Repeated measures, within-between interaction			
Analysis:	A priori: Compute required sample size		
Input:	Effect size f	=	0.175
	α err prob	=	0.05
	Power (1−β err prob)	=	0.8
	Number of groups	=	2
	Number of measurements	=	2
	Corr among rep measures	=	0.5
	Nonsphericity correction ϵ	=	1
Output:	Noncentrality parameter λ	=	8.3300000
	Critical F	=	3.9862695
	Numerator df	=	1.0000000
	Denominator df	=	66.0000000
	Total sample size	=	68
	Actual power	=	0.8116461

2 x 2 x 2 interaction effect involving 2 independent measures variables and 1 repeated measures variable (All one factor with 2 levels)

F tests – ANOVA: Repeated measures, within-between interaction			
Analysis:	A priori: Compute required sample size		
Input:	Effect size f	=	0.175
	α err prob	=	0.05
	Power (1−β err prob)	=	0.8
	Number of groups	=	4
	Number of measurements	=	2
	Corr among rep measures	=	0.5
	Nonsphericity correction ϵ	=	1
Output:	Noncentrality parameter λ	=	11.7600000
	Critical F	=	2.7035940
	Numerator df	=	3.0000000
	Denominator df	=	92.0000000
	Total sample size	=	96
	Actual power	=	0.8143056

2 x 2 x 2 interaction effect involving 1 independent measures variable and 2 repeated measures variables (All one factor with 2 levels)

F tests – ANOVA: Repeated measures, within–between interaction			
Analysis:	A priori: Compute required sample size		
Input:	Effect size f	=	0.175
	α err prob	=	0.05
	Power (1–β err prob)	=	0.8
	Number of groups	=	2
	Number of measurements	=	4
	Corr among rep measures	=	0.5
	Nonsphericity correction ϵ	=	1
Output:	Noncentrality parameter λ	=	11.2700000
	Critical F	=	2.6732178
	Numerator df	=	3.0000000
	Denominator df	=	132
	Total sample size	=	46
	Actual power	=	0.8015327

2 x 2 x 2 x 2 interaction effect involving 2 independent measures variables and 2 repeated measures variables (All one factor with 2 levels)

F tests – ANOVA: Repeated measures, within–between interaction			
Analysis:	A priori: Compute required sample size		
Input:	Effect size f	=	0.175
	α err prob	=	0.05
	Power (1–β err prob)	=	0.8
	Number of groups	=	4
	Number of measurements	=	4
	Corr among rep measures	=	0.5
	Nonsphericity correction ϵ	=	1
Output:	Noncentrality parameter λ	=	16.6600000
	Critical F	=	1.9289038
	Numerator df	=	9.0000000
	Denominator df	=	192
	Total sample size	=	68
	Actual power	=	0.8091441